

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 784 926**

51 Int. Cl.:

**C12N 15/79** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.12.2008** **E 16168059 (0)**

97 Fecha y número de publicación de la concesión europea: **29.01.2020** **EP 3075858**

54 Título: **Vector de expresión de mamífero**

30 Prioridad:

**21.12.2007 EP 07150339**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**02.10.2020**

73 Titular/es:

**NOVARTIS AG (100.0%)  
Lichtstrasse 35  
4056 Basel , CH**

72 Inventor/es:

**KNOPF, HANS-PETER y  
WILMS, BURKHARD**

74 Agente/Representante:

**CARVAJAL Y URQUIJO, Isabel**

**ES 2 784 926 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Vector de expresión de mamífero

5 La presente invención se refiere a un vector de expresión de mamífero para expresar un polipéptido de interés así como a métodos para expresar un polipéptido de interés en una célula hospedadora de mamífero utilizando un vector respectivo y células hospedadoras que comprenden dicho vector.

10 La capacidad para clonar y expresar péptidos y proteínas recombinantes en cantidades grandes ha cobrado cada vez mayor importancia en los últimos años. La capacidad para purificar niveles altos de proteínas es importante en el campo biotecnológico y farmacéutico humano, por ejemplo, para producir fármacos proteínicos así como en el ámbito de la investigación básica, por ejemplo, para cristalizar proteínas con el fin de poder determinar la estructura tridimensional. Las proteínas que de otro modo son difíciles de obtener en cantidad se pueden sobreexpresar en una célula hospedadora, y subsiguientemente aislar y purificar.

15 La elección de un sistema de expresión para la producción de proteínas recombinantes depende de muchos factores, que incluyen las características de crecimiento celular, los niveles de expresión, la expresión intracelular y extracelular, las modificaciones postraduccionales y la actividad biológica de la proteína de interés, así como cuestiones normativas y consideraciones económicas en la producción de proteínas terapéuticas. Las principales ventajas de las células de mamífero frente a otros sistemas de expresión tales como bacterias o levadura son la capacidad para llevar a cabo un plegamiento proteínico correcto, *N*-glucosilación compleja y *O*-glucosilación auténtica, así como un amplio espectro de otras modificaciones postraduccionales. Debido a las ventajas descritas, las células de mamífero son en la actualidad el sistema de expresión preferido para producir proteínas terapéuticas complejas tales como anticuerpos monoclonales. El primer paso en la generación de un línea celular recombinante es la construcción de un vector de expresión. El vector de expresión es el elemento clave para dirigir la expresión del gen heterólogo en la célula hospedadora y para proporcionar marcadores de selección con el fin de generar la línea celular recombinante. Los elementos esenciales de los vectores de expresión de mamífero normalmente incluyen un promotor constitutivo o inducible que dispone de actividad transcripcional sólida; señales de traducción y procesamiento de ARNm optimizado que normalmente incluyen una secuencia de Kozak, un codón de terminación de la traducción, señales de escisión y poliadenilación de ARNm, así como señales de corte y empalme de ARNm; un terminador de la transcripción; marcadores de selección para la preparación de líneas celulares estables y para la amplificación de genes; un origen procariótico de replicación y marcadores de selección para la propagación de vectores en bacterias.

35 En los últimos años, el objetivo principal de la investigación se centraba en el diseño de vectores mejorados para la expresión de genes en células de mamífero. A pesar de la multitud de vectores disponibles, sin embargo, la producción consistente de polipéptidos/proteínas en células de mamífero sigue representando un reto. Varios factores pueden afectar a la expresión recombinante en células de mamífero, que incluyen la potencia del promotor, el contexto de la región de iniciación de la traducción y no traducida 5', la eficiencia de la región no traducida 3', para poliadenilar y terminar la transcripción, el sitio de inserción del gen recombinante integrado aleatoriamente en el cromosoma hospedador y el número de copias integradas del gen que se esté expresando. Lo más habitual es que los incrementos en el número de copias del gen se consigan mediante amplificación génica utilizando líneas celulares deficientes en una enzima tal como la dihidrofolato reductasa (DHFR) o glutamina sintetasa (GS) junto con vectores de expresión que contienen genes que codifican estas enzimas y agentes tales como metotrexato (MTX), que inhibe DHFR, y metionina sulfoxamina (MSX) que inhibe GS. Utilizando vectores de expresión que contienen el gen recombinante bajo el control de un promotor potente y genes que codifican DHFR o GS, se obtienen primero transfectantes DHFR<sup>+</sup> o GS<sup>+</sup>, respectivamente, y la amplificación génica se consigue entonces cultivando los transfectantes en concentraciones progresivamente mayores de MTX o MSX.

50 El documento WO2004/050879 presenta un vector de expresión que comprende un casete de expresión bicistrónico que comprende el promotor de ubiquitina/S27a de hámster para expresar el gen de interés y un gen indicador tal como GFP. En el vector circular, el gen dhfr está situado en 5' respecto al casete de expresión que comprende el gen que codifica el polipéptido de interés y el gen indicador GFP.

55 El documento WO01/04306 presenta un vector de expresión que comprende un gen marcador seleccionable amplificable tal como dhfr y un gen indicador, concretamente, GFP. Presenta un sistema de expresión bicistrónico en el que la expresión del gen que codifica la proteína de interés está ya sea ligada operativamente a la expresión del marcador seleccionable amplificable o al gen indicador GFP.

60 Lin *et al.*, «Expression efficiency of the human thrombomodulin-encoding gene in various vector and host systems» en *Gene*, 147 (1994) 287-292 analiza el comportamiento de cinco vectores de expresión de mamífero que expresan trombomodulina humana recombinante en células hospedadoras diferentes.

El documento US 5.017.478 presenta plásmidos de expresión eucarióticos que comprenden un casete de gen seleccionable tal como dhfr y un casete génico que codifica el producto de interés, donde el casete de gen seleccionable y el casete génico para expresar la proteína de interés están dispuestos adyacentes uno respecto al otro en orientación transcripcional divergente y opuesta. El casete de gen seleccionable está situado en 5' del casete génico para expresar la proteína de interés. Un casete de gen marcador seleccionable adicional que codifica, p. ej., neo o higromicina, puede estar situado en 3' del casete génico que codifica el producto de interés. A efectos de comparación, se diseñó un vector de expresión pPA206, entre otros, en el que el casete de gen seleccionable que codifica dhfr, el casete génico que codifica el producto de interés y el casete de gen seleccionable que codifica neo están orientados en la misma dirección.

5

Es el objetivo de la presente invención proporcionar un vector de expresión mejorado así como un sistema de expresión para expresar un polipéptido de interés en una célula de mamífero.

10

La presente invención puede estar definida por las reivindicaciones adjuntas.

Por consiguiente, de acuerdo con la reivindicación 1, se proporciona un ácido nucleico de vector linealizado adecuado para expresar al menos un polipéptido de interés en una célula de mamífero, que comprende

15

(a) al menos un casete de expresión (POI) adecuado para expresar un polipéptido de interés, donde dicho casete de expresión comprende el polinucleótido que codifica el polipéptido de interés y comprende al menos un promotor o elemento promotor/potenciador;

20

(b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero, donde dicho gen marcador seleccionable de mamífero es un gen de resistencia a antibiótico y donde dicho casete de expresión comprende al menos un promotor o elemento promotor/potenciador;

25

(c) un casete de expresión (MASM) que comprende un gen marcador seleccionable, amplificable de mamífero, donde dicho casete de expresión (MASM) comprende al menos un promotor o elemento promotor/potenciador; y

30

(d) un casete de expresión (PSM) que comprende un gen marcador seleccionable procariótico;

30

donde el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está situado en 3' respecto al casete de expresión (POI) y el casete de expresión (PSM) está situado en 3' del casete de expresión (MSM), y donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3'.

35

También se da a conocer un ácido nucleico de vector para expresar al menos un polipéptido de interés en una célula de mamífero, que comprende

40

(a) al menos un casete de expresión (POI) para expresar un polipéptido de interés;

40

(b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero;

45

(c) un casete de expresión (MASM) que comprende un gen marcador seleccionable, amplificable de mamífero;

45

donde el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está situado en 3' respecto al casete de expresión (POI), y donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3'.

50

Un «ácido nucleico de vector» de acuerdo con la presente invención es un polinucleótido que porta al menos un fragmento de ácido nucleico exógeno. Un ácido nucleico de vector funciona como un «portador molecular», suministrando fragmentos de ácidos nucleicos respectivamente polinucleótidos a una célula hospedadora. Comprende al menos un casete de expresión que comprende secuencias reguladoras y codificantes. Un casete de expresión permite la expresión apropiada de un polinucleótido incorporado. Los casetes de expresión decisivos para la presente invención se describirán subsiguientemente en más detalle. Se pueden insertar polinucleótidos exógenos, p. ej., que codifican el polipéptido de interés, en los casetes de expresión del ácido nucleico de vector con el fin de que se exprese. El ácido nucleico de vector de acuerdo con la presente invención está presente en forma linealizada. Dicho sitio puede ser, p. ej., un sitio de clonación múltiple (MCS, por sus siglas en inglés).

55

60

El casete de expresión (POI) define el casete de expresión para expresar un polipéptido de interés (*polypeptide of interest*). Dicho casete de expresión (POI) comprende ya sea el polinucleótido que codifica el polipéptido de interés o comprende un sitio adecuado para insertar un polinucleótido respectivo que codifica el polipéptido de interés.

Un «polipéptido» se refiere a una molécula que comprende un polímero de aminoácidos unidos entre sí por uno o más enlaces peptídicos. Los polipéptidos incluyen polipéptidos de cualquier longitud, que incluyen proteínas (por ejemplo, que tienen más de 50 aminoácidos) y péptidos (por ejemplo, que tienen 2-49 aminoácidos). Los polipéptidos incluyen proteínas y/o péptidos de cualquier actividad o bioactividad. Más adelante se exponen ejemplos adecuados.

El casete de expresión (MSM) define el casete de expresión que comprende un gen marcador seleccionable de mamífero (*mammalian selectable marker gene*) que es un gen de resistencia a antibiótico. Los genes marcadores seleccionables de mamífero permiten la selección de células hospedadoras de mamífero que comprenden dichos genes y, por tanto, de células hospedadoras de mamífero que comprenden el vector. Más adelante se describen ejemplos adecuados en detalle.

El casete de expresión (MASM) define el casete de expresión que comprende un gen marcador seleccionable, amplificable de mamífero (*mammalian amplifiable, selectable marker gene*). Los genes marcadores seleccionables, amplificables de mamífero permiten la selección de células hospedadoras de mamífero que contienen el vector así como la amplificación génica. Más adelante se describen ejemplos adecuados en detalle.

Los términos «5'» y «3'» es una convención utilizada para describir características de una secuencia de ácido nucleico relacionadas con ya sea la posición de elementos genéticos y/o la dirección de eventos (5' a 3'), tales como, p. ej., la transcripción por parte de la ARN·polimerasa o traducción por parte del ribosoma que avanza en dirección 5' a 3'. Son sinónimos secuencia arriba (5') y secuencia abajo (3'). Convencionalmente, las secuencias de ADN, los mapas genéticos, tarjetas de vectores y secuencias de ARN se dibujan con 5' a 3' de izquierda a derecha o la dirección 5' a 3' se indica con flechas, donde la punta de las flechas apunta en la dirección 3'. Por consiguiente, 5' (secuencia arriba) indica elementos genéticos colocados hacia el lado izquierdo, y 3' (secuencia abajo) indica elementos genéticos colocados hacia el lado derecho, cuando se sigue esta convención.

La disposición y la orientación de los casetes de expresión presentes en el vector de la invención es importante. De acuerdo con las enseñanzas de la presente invención, el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM). Por consiguiente, el casete de expresión (MASM) está situado en 5' adyacente al casete de expresión (POI) y en estrecha proximidad con él. Evidentemente, los casetes de expresión (MASM) y (POI) pueden estar separados por secuencias de esqueleto del vector. Sin embargo, preferentemente, no hay ningún otro casete de expresión situado entre el casete de expresión (MASM) y el casete de expresión (POI). El casete de expresión (MSM) está situado en 3' respecto al casete de expresión (POI). Puede haber casetes de expresión adicionales insertados entre los casetes de expresión (POI) y (MSM), tales como, p. ej., un casete de expresión adicional (POI') para expresar un polipéptido de interés adicional (descrito en más detalle más adelante). Los casetes de expresión (MASM), (POI) y (MSM) están todos dispuestos en la misma orientación 5' a 3'. Los inventores descubrieron que esta configuración de ácido nucleico de vector particular permite la generación rápida de líneas celulares de alto rendimiento.

El polinucleótido que codifica el polipéptido de interés se puede incorporar en el casete de expresión (POI) utilizando métodos de clonación apropiados, por ejemplo, utilizando enzimas de restricción con el fin de insertar el polinucleótido que codifica el polipéptido de interés en el casete de expresión (POI). Con este propósito, el casete de expresión (POI) puede comprender, p. ej., un sitio de clonación múltiple (MCS) que se puede, p. ej., utilizar en todos los marcos de lectura. Más adelante se describen sitios MCS adecuados en detalle. Se puede proporcionar, p. ej., un ácido nucleico de vector «vacío» respectivo a los clientes, que entonces insertan su polinucleótido de interés específico que desean expresar en el casete de expresión (POI). El casete de expresión (POI) también puede comprender un polinucleótido de reemplazo o una secuencia de ácido nucleico de relleno, que se puede extirpar y reemplazar con el polinucleótido que codifica el polipéptido de interés. La presente invención proporciona un ácido nucleico de vector tal como se ha descrito anteriormente, que comprende un casete de expresión (POI) que comprende el polinucleótido que codifica el polipéptido de interés. El ácido nucleico de vector de expresión final se puede transfectar para su expresión en la célula hospedadora.

Un polinucleótido es un polímero de nucleótidos que normalmente están unidos de una desoxirribosa o ribosa a otra. El término «polinucleótido» no comprende ninguna restricción de tamaño y también abarca polinucleótidos que comprenden modificaciones, en particular nucleótidos modificados.

En un ácido nucleico de vector circular que se da a conocer en la presente, el casete de expresión (MSM) está dispuesto en 3' del casete de expresión (POI) y el casete de expresión (MASM) está dispuesto en 3' del casete de expresión (MSM). Una descripción alternativa de un vector circular de este tipo es un ácido nucleico de vector circular para expresar al menos un polipéptido de interés en una célula de mamífero, que comprende

- (a) al menos un casete de expresión (POI) para expresar un polipéptido de interés;
- (b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero;
- (c) un casete de expresión (MASM) que comprende un gen marcador seleccionable, amplificable de mamífero;

donde el casete de expresión (MSM) está dispuesto en 3' del casete de expresión (POI) y el casete de expresión (MASM) está dispuesto en 3' del casete de expresión (MSM), y donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3'.

5

Un ácido nucleico de vector se puede transfectar en la célula hospedadora en su forma circular. Las moléculas de vector superhelicoidales se convertirán normalmente en moléculas lineales dentro del núcleo debido a la actividad de endo- y exonucleasas. La linealización del ácido nucleico de vector antes de la transfección suele mejorar la eficiencia de una transfección estable. Esto también ya que el punto de linealización se puede controlar si el vector se linealiza antes de la transfección.

10

Un vector de expresión circular puede comprender un sitio de restricción predefinido, que se puede utilizar para la linealización del ácido nucleico de vector antes de la transfección. La ubicación inteligente de dicho sitio de restricción para la linealización es importante, porque dicho sitio de restricción determina el lugar en el que el ácido nucleico de vector se abre/linealiza y, por tanto, determina el orden/la disposición de los casetes de expresión cuando el constructo se integra en el genoma de la célula de mamífero.

15

Por consiguiente, un ácido nucleico de vector circular puede comprender un sitio de restricción para la linealización para linealizar el vector, donde dicho sitio de restricción para la linealización está situado entre los casetes de expresión (MSM) y (MASM). Preferentemente, dicho sitio de restricción para la linealización es único y solamente está presente una vez en el ácido nucleico de vector de expresión. P. ej., se puede utilizar un sitio de restricción para la linealización que sea reconocido por una enzima de restricción con una frecuencia de corte baja con el fin de fomentar que el vector se escinda solamente en el sitio de restricción para la linealización pero no (o solamente rara vez), p. ej., dentro del casete o casetes de expresión o el esqueleto del vector. Esto se puede propiciar, p. ej., proporcionando un sitio de restricción para una enzima de restricción con una secuencia de reconocimiento de más de seis pares de bases o que reconoce secuencias que están infrarrepresentadas en el ADN cromosómico. Un ejemplo adecuado es la enzima *SwaI* y el vector puede incorporar, por lo tanto, un sitio de reconocimiento de *SwaI* como único sitio de restricción para la linealización. En el caso de que dicho sitio de restricción para la linealización esté presente más de una vez en la secuencia de ácido nucleico de vector (que incluye los polinucleótidos que codifican el polipéptido de interés), o en el caso en que se utilice una enzima de restricción que corta varias veces en la secuencia de ácido nucleico de vector, también entra dentro del alcance de la presente invención, p. ej., alterar/mutar los sitios de restricción aparte del sitio de restricción para la linealización que está situado entre los casetes de expresión (MSM) y (MASM), con el fin de eliminar esos sitios de restricción adicionales y para obtener un único o al menos infrecuente sitio de restricción para la linealización.

20

25

30

35

En el caso de que el vector se utilice como vector de expresión estándar diseñado, p. ej., como herramienta para la expresión de varios polipéptidos diferentes, es ventajoso proporcionar un sitio de restricción para la linealización que comprenda múltiples sitios de reconocimiento para enzimas que tengan una frecuencia de corte baja. Las enzimas de restricción elegidas para la linealización preferentemente no deben cortar dentro de los casetes de expresión para los marcadores seleccionables u otras secuencias de esqueleto del vector con el fin de garantizar que la enzima corte solamente una vez para la linealización adecuada del vector. Al proporcionar un sitio de restricción para la linealización que comprende múltiples sitios de reconocimiento para enzimas de restricción que tienen una frecuencia de corte baja, el usuario puede elegir una enzima de restricción adecuada para la linealización entre las opciones proporcionadas con el fin de evitar de forma segura la restricción dentro del polinucleótido que codifica el polipéptido de interés. Sin embargo, tal como se ha expuesto anteriormente, se pueden mutar sitios de restricción adicionales o se podría realizar una digestión de restricción parcial.

40

45

La colocación del sitio de restricción para la linealización entre el casete de expresión (MSM) y el casete de expresión (MASM) tiene el efecto de que el casete de expresión (POI) (y casetes de expresión adicionales para expresar los polipéptidos de interés - si están presentes) está flanqueado en 5' por el casete de expresión (MASM). El casete de expresión (MSM) está situado en 3' del casete de expresión (POI) tras la linealización. De este modo, los casetes de expresión (MSM) y (MASM) están separados tras la linealización del ácido nucleico de vector circular. Como hay un casete de expresión (PSM) para un marcador de selección bacteriano presente en el ácido nucleico de vector linealizado reivindicado (véase también más adelante), el sitio de restricción para la linealización está preferentemente ubicado entre los casetes de expresión (PSM) y (MASM). Esto tiene el efecto de que el gen marcador de selección bacteriano está en 3' y, por tanto, «fuera» de las partes «de mamífero» del ácido nucleico de vector linealizado. Esta disposición es favorable ya que los genes bacterianos no son presumiblemente ventajosos para la expresión en mamíferos ya que las secuencias bacterianas pueden provocar una mayor metilación u otros efectos de silenciamiento en las células de mamífero.

50

55

Los ejemplos no limitantes para genes marcadores seleccionables de mamífero que pueden estar comprendidos en el casete de expresión (MSM) incluyen genes de resistencia a antibiótico, p. ej., que confieren resistencia a G418; higromicina (*hyg* o *hph*, comercializada por Life Technologies, Inc. Gaithersboro, Md.); neomicina (*neo*, comercializada por Life Technologies, Inc. Gaithersboro, Md.); zeocina (*Sh Ble*, comercializada por Pharmingen, San Diego Calif.); puromicina

60

(pac, puromicina-*N*-acetil-transferasa, comercializada por Clontech, Palo Alto Calif.), ouabaína (oua, comercializada por Pharmingen) y blasticidina (comercializada por Invitrogen). Los respectivos genes marcadores seleccionables de mamífero son muy conocidos y permiten la selección de células hospedadoras de mamífero que comprenden dichos genes y, por tanto, de células hospedadoras que comprenden el vector.

5 El término «gen», tal como se utiliza en la presente, también se refiere a un polinucleótido natural o sintético que codifica una variante funcional del marcador seleccionable que proporciona la resistencia deseada. Por ende, también se abarcan versiones truncadas o mutadas de un gen natural o polipéptidos sintéticos siempre que proporcionen la resistencia deseada. De acuerdo con una realización preferida, dicho casete de expresión (MSM) comprende un gen que codifica una neomicina·fosfotransferasa (I o II) enzimáticamente funcional. Esta realización funciona bien en combinación con el uso de un gen que codifica una DHFR enzimáticamente funcional como gen marcador seleccionable amplificable.

15 Los genes marcadores seleccionables de mamífero, amplificables permiten la selección de células hospedadoras de mamífero que contienen el vector así como la amplificación génica. Un ejemplo no limitante para un gen marcador seleccionable de mamífero, amplificable es el gen de la dihidrofolato·reductasa (DHFR). Otros sistemas que se utilizan en la actualidad son, entre otros, el sistema de la glutamina·sintetasa (gs) (Bebbington *et al.*, 1992) y el sistema de selección dependiente de histidinol (Hartmann y Mulligan, 1988). Estos marcadores amplificables también son marcadores seleccionables y, por tanto, se pueden utilizar para seleccionar aquellas células que obtuvieron el vector. La DHFR y la glutamina·sintetasa proporcionan buenos resultados. En ambos casos, la selección normalmente se produce en ausencia del metabolito apropiado (hipoxantina y timidina en el caso de la DHFR, glutamina en el caso de la GS), lo cual evita el crecimiento de células no transformadas. Con sistemas amplificables tales como el sistema DHFR, la expresión de una proteína recombinante se puede incrementar exponiendo las células a ciertos agentes que fomentan la amplificación génica tales como antifolatos (p. ej., metotrexato (MTX)) en el caso del sistema DHFR. Un inhibidor adecuado para GS que fomenta la amplificación génica es la metionina sulfoximina (MSX). La exposición a MSX también provoca la amplificación génica. De acuerdo con una realización, dicho casete de expresión (MASM) comprende un gen que codifica una glutamina·sintetasa (GS) enzimáticamente funcional o una dihidrofolato·reductasa (DHFR).

20 De acuerdo con una realización, dicho casete de expresión (MSM) comprende un gen que codifica una neomicina·fosfotransferasa enzimáticamente funcional y dicho casete de expresión (MASM) comprende un gen que codifica una dihidrofolato·reductasa (DHFR) enzimáticamente funcional.

25 El vector puede comprender al menos un casete expresión adicional (POI') para expresar un polipéptido de interés adicional. Dicho casete de expresión adicional (POI') está situado entre el casete de expresión (POI) y el casete de expresión (MSM). Dicho casete de expresión (POI') está dispuesto en la misma orientación 5' a 3' que los casetes de expresión (POI) y (MSM). De acuerdo con una realización, comprende el polinucleótido que codifica el polipéptido de interés adicional.

30 Por consiguiente, el ácido nucleico de vector linealizado de acuerdo con la presente invención puede comprender más de un casete de expresión para expresar polipéptidos de interés.

35 Por lo tanto, también es posible que varios casetes de expresión ((POI), (POI'), (POI'') etc.) para expresar diferentes polipéptidos de interés estén dispuestos en el ácido nucleico de vector de expresión linealizado de acuerdo con la presente invención. Estos casetes de expresión están flanqueados en 5' por el casete de expresión (MASM) y en 3' por el casete de expresión (MSM). Por ende, la presente invención también proporciona un ácido nucleico de vector linealizado que comprende más de un casete de expresión que codifica, p. ej., subunidades de proteínas dimericas o multiméricas de orden superior. Los casetes de expresión que codifican diferentes subunidades de una proteína multimérica, incorporada cada una en un casete de expresión diferente, pueden estar ubicados adyacentes entre sí. Para proteínas multiméricas codificadas por al menos dos genes distintos (por ejemplo, cadenas ligera y pesada de inmunoglobulina o fragmentos funcionales de estas tales como al menos las regiones variables de las cadenas ligera y pesada de inmunoglobulina), los polinucleótidos que codifican las subunidades deseadas del polipéptido de interés están insertados en los casetes de expresión (POI) y (POI'). Una realización respectiva que utiliza al menos dos casetes de expresión (POI) y (POI') para expresar polipéptidos de interés es particularmente ventajosa para expresar moléculas de inmunoglobulina tales como anticuerpos o fragmentos funcionales de estos. Por consiguiente, se proporciona un ácido nucleico de vector para expresar una molécula de inmunoglobulina que comprende en cada casete de expresión (POI) y (POI') un polinucleótido que codifica ya sea una cadena ligera o una pesada de una molécula de inmunoglobulina o fragmentos de esta, donde cada casete de expresión (POI) y (POI') comprende uno de dichos polinucleótidos. Por consiguiente, el casete de expresión (POI) puede comprender ya sea el polinucleótido para expresar la cadena ligera, o el polinucleótido para expresar la cadena pesada de la molécula de inmunoglobulina.

40 De acuerdo con una realización preferida, el casete de expresión (POI) comprende el polinucleótido que codifica al menos parte de la cadena ligera de dicha molécula de inmunoglobulina o un fragmento funcional de esta y el casete de expresión (POI') comprende un polinucleótido que codifica al menos parte de la cadena pesada de dicha molécula de

5 inmunoglobulina o un fragmento funcional de esta. Colocar el casete de expresión para la cadena ligera en 5' respecto al casete de expresión de la cadena pesada resultó ser beneficioso en lo que respecta a la tasa de expresión de las moléculas de inmunoglobulina. De acuerdo con una realización, el ácido nucleico de vector de expresión se diseña de tal modo que el casete o casetes de expresión ya comprenda un polinucleótido que codifica al menos parte de las regiones constantes de una molécula de inmunoglobulina. Los fragmentos del polinucleótido que codifican las partes variables de las moléculas de inmunoglobulina pueden ser insertados entonces por el usuario/cliente en los casetes de expresión utilizando estrategias de clonación apropiadas con el fin de obtener el vector de expresión final.

10 Los casetes de expresión presentes en el vector de expresión linealizado de acuerdo con la presente invención están diseñados de tal modo que permitan la expresión de los polinucleótidos/genes incorporados en células de mamífero. A este efecto, los casetes de expresión normalmente comprenden las secuencias reguladoras necesarias, tales como un promotor y/o una secuencia de terminación de la transcripción tal como un sitio de poli A.

15 Los vectores utilizados para expresar polipéptidos de interés normalmente contienen elementos de control de la transcripción adecuados para dirigir la transcripción tales como, p. ej., promotores, potenciadores, señales de poliadenilación, señales de pausa o terminación de la transcripción como elementos de un casete de expresión. Para la expresión apropiada de los polipéptidos, preferentemente se incluyen en el vector elementos de control de la traducción adecuados, tales como, p. ej., regiones no traducidas 5' que dan lugar a estructuras de caperuza 5' adecuadas para captar ribosomas y codones de parada para terminar el proceso de traducción. El polinucleótido que actúa como los genes marcadores seleccionables así como el polinucleótido que codifica el polipéptido de interés, se transcriben bajo el control de elementos de transcripción presentes en promotores apropiados. Los transcritos resultantes de los genes marcadores seleccionables y el del polipéptido de interés albergan elementos de traducción funcionales que facilitan niveles sustanciales de expresión proteínica (es decir, traducción) y una terminación de la traducción apropiada. Una unidad de expresión funcional, capaz de dirigir de forma apropiada la expresión de un polinucleótido incorporado también se denomina «casete de expresión» en la presente. Preferentemente, comprende una región UTR 3'.

20 Los casetes de expresión comprenden al menos un promotor y/o elemento promotor/potenciador. Aunque las limitaciones físicas entre estos dos elementos de control no están siempre claras, el término «promotor» normalmente se refiere a un sitio en la molécula de ácido nucleico al que se une una ARN·polimerasa y/o cualesquiera factores asociados y en el cual se inicia la transcripción. Los potenciadores potencian la actividad del promotor, temporalmente así como espacialmente. Muchos promotores son activos transcripcionalmente en una gran diversidad de tipos celulares. Los promotores se pueden dividir en dos clases, los que funcionan de forma constitutiva y los que están regulados por inducción o desrepresión. Los promotores utilizados para la producción en niveles altos de proteínas en células de mamífero deben ser potentes y preferentemente activos en una gran diversidad de tipos celulares para permitir una evaluación cualitativa y cuantitativa del polipéptido recombinante. El promotor se puede seleccionar del grupo que consiste en un promotor de SV40, un promotor de CMV, un promotor de EF1alfa, un promotor de RSV, un promotor de BROAD3, un promotor de rosa26 murino, un promotor de pCEFL y un promotor de  $\beta$ -actina. Los promotores constitutivos potentes que dirigen la expresión en muchos tipos celulares incluyen, pero no se limitan a, el promotor tardío principal de adenovirus, el promotor temprano inmediato de citomegalovirus humano, el promotor de SV40 y del virus del sarcoma de Rous, y el promotor de la 3-fosfoglicerato·cinasa murina y EF1a. Se consiguen buenos resultados con el vector de expresión de la presente invención cuando el promotor y/o potenciador se obtiene ya sea a partir de CMV y/o SV40.

45 De acuerdo con una realización, el casete o casetes de expresión para expresar el polipéptido o polipéptidos de interés comprenden un promotor y/o potenciador más potente que los casetes de expresión para expresar los marcadores seleccionables. Esta disposición tiene el efecto de que se genera más transcrito para el polipéptido de interés que para los marcadores de selección. Es ventajoso que la producción del polipéptido de interés que se secreta sea dominante frente a la producción de los marcadores de selección, dado que la capacidad celular individual para producir proteínas heterólogas no es ilimitada y, por tanto, se debe centrar en el polipéptido de interés.

50 De acuerdo con una realización, los casetes de expresión (POI) y (POI') (si está presente) que se utiliza/n para expresar el polipéptido de interés comprenden un promotor/potenciador de CMV. Más adelante se describen ejemplos específicos en detalle. Los casetes de expresión (MSM) y (MASM), que expresan preferentemente la DHFR y los genes marcadores de neomicina, comprenden un promotor de SV40 o un promotor/potenciador de SV40. Se sabe que el promotor de CMV es uno de los promotores más potentes disponibles para la expresión en mamíferos y conduce a una tasa de expresión muy buena. Se considera que proporciona significativamente más transcrito que el promotor de SV40.

60 Además, los casetes de expresión pueden comprender un sitio de terminación de la transcripción apropiado. Esto, como la transcripción continuada desde un promotor secuencia arriba a través de una segunda unidad de transcripción puede inhibir la función del promotor secuencia abajo, un fenómeno conocido como la oclusión del promotor o interferencia transcripcional. Este evento se ha descrito tanto en procariontes como eucariotes. La ubicación apropiada de señales de terminación de la transcripción entre dos unidades de transcripción puede prevenir la oclusión del promotor. Los sitios de

terminación de la transcripción están bien caracterizados y se ha demostrado que su incorporación en vectores de expresión tiene múltiples efectos beneficiosos sobre la expresión génica.

La mayoría de los ARNm nacientes eucarióticos poseen un cola de poli A en su extremo 3' que se añade durante un proceso complejo que implica la escisión del transcrito primario y una reacción de poliadenilación acoplada. La cola de poliA es ventajosa para la estabilidad y transferibilidad del ARNm. Por ende, los casetes de expresión del vector linealizado de acuerdo con la presente invención normalmente comprenden un sitio de poliadenilación. Existen varias señales de poliA eficientes que se pueden utilizar en vectores de expresión de mamífero, que incluyen las derivadas de la hormona del crecimiento bovina (bgh, por sus siglas en inglés), beta-globina de ratón, la unidad de transcripción temprana de SV40 y el gen de la timidina·cinasa del virus del herpes simple. Sin embargo, también se conocen sitios de poliadenilación sintéticos (véase, p. ej., el vector de expresión pCI-neo de Promega que se basa en Levitt *et al.*, 1989, *Genes Dev.* 3, (7): 1019-1025). El sitio de poliadenilación se puede seleccionar del grupo que consiste en el sitio de poliA de SV40, tal como el sitio de poli-A tardío y temprano de SV40 (véase, p. ej., el plásmido pSV2-DHFR tal como se describe en Subramani *et al.*, 1981, *Mol.Cell. Biol.* 854-864), un sitio de poliA sintético (véase, p. ej., el vector de expresión pCI-neo de Promega que se basa en Levitt *et al.*, 1989, *Genes Dev.* 3, (7): 1019-1025) y un sitio de poliA de bgh (hormona del crecimiento bovina).

Los casetes de expresión pueden comprender un potenciador (véase anteriormente) y/o un intrón. De acuerdo con una realización, el casete o casetes de expresión para expresar el polipéptido de interés comprenden un intrón. La mayoría de los genes de eucariotas superiores contienen intrones que se eliminan durante el procesamiento del ARN. Se observa que los constructos genómicos se expresan de forma más eficiente en sistemas transgénicos que constructos idénticos que carecen de intrones. Normalmente, los intrones están ubicados en el extremo 5' del marco abierto de lectura. Por consiguiente, un intrón puede estar comprendido en el casete o casetes de expresión para expresar el polipéptido o polipéptidos de interés con el fin de incrementar la tasa de expresión. Dicho intrón puede estar situado entre el promotor y o uno o más elementos promotores/potenciadores y el extremo 5' del marco abierto de lectura del polinucleótido que se desea expresar. Por ende, se proporciona un ácido nucleico de vector, donde al menos el casete de expresión (POI) comprende un intrón que está dispuesto entre el promotor y el condón de iniciación del polinucleótido para expresar el polipéptido de interés. Se conocen varios intrones adecuados en el estado de la técnica que se pueden utilizar junto con la presente invención.

De acuerdo con una realización, el intrón utilizado en los casetes de expresión para expresar los polipéptidos de interés, es un intrón sintético tal como el intrón SIS o el RK. El intrón RK es un intrón sintético potente que está ubicado preferentemente antes del codón de iniciación ATG del gen de interés. El intrón RK consiste en el sitio de corte y empalme donador intrónico del promotor de CMV y el sitio de corte y empalme aceptor de la región variable de cadena pesada de IgG (véanse, p. ej., Eaton *et al.*, 1986, *Biochemistry* 25, 8343-8347, Neuberger *et al.*, 1983, *EMBO J.* 2(8), 1373-1378; se puede obtener a partir del vector pRK-5 (BD Pharmingen)).

Además, se ha descubierto sorprendentemente que la ubicación de un intrón en el extremo 3' del marco abierto de lectura del gen DHFR tiene efectos ventajosos sobre la tasa de expresión/amplificación del constructo. El intrón utilizado en el casete de expresión de DHFR da lugar a una variante no funcional, más pequeña del gen DHFR (Grillari *et al.*, 2001, *J. Biotechnol.* 87, 59-65). De este modo, el nivel de expresión del gen DHFR se reduce. Esto conduce a una mayor sensibilidad a MTX y condiciones de selección más restrictivas. Por consiguiente, se proporciona un ácido nucleico de vector, donde el casete de expresión (MASM) comprende un intrón que está situado en 3' del gen marcador seleccionable amplificable. Se puede obtener un intrón adecuado a partir del vector pSV2-DHFR (véase, p. ej., anteriormente).

El vector reivindicado comprende al menos un casete de expresión adicional (PSM) que comprende un gen marcador seleccionable procarriótico. Dicho casete de expresión (PSM) está situado entre los casetes de expresión (MSM) y (MASM). Dicho marcador seleccionable puede proporcionar una resistencia a antibiótico tal como, p. ej., ampicilina, kanamicina, tetraciclina y/o cloranfenicol. Dicho casete de expresión (PSM) está dispuesto preferentemente en la misma orientación 5' a 3' que los otros casetes de expresión (POI), (MSM) y (MASM).

El gen DHFR funciona como un marcador y un gen de amplificación génica. La selección puede ocurrir mediante el cultivo de las células en ausencia de los metabolitos apropiados (hipoxantina y timidina) previniendo de este modo el crecimiento de células no transformadas. Con el sistema DHFR, se puede incrementar la expresión de los polipéptidos de interés al exponer las células a antifolatos tales como, por ejemplo, metotrexato (MTX), un fármaco que bloquea la actividad de DHFR. Después de un tiempo determinado de exposición a MTX, la mayoría de las células mueren, pero normalmente sobreviven un pequeño número de células que sobreproducen DHFR. Al tratarlas con MTX, en concentraciones crecientes, a menudo las células supervivientes pueden contener hasta de varios cientos a unos pocos miles de copias del vector integrado embebido en cromosomas que suelen estar elongados. De la forma más respectiva, las células amplificadas producen más proteína recombinante que las células sin amplificar.

En la técnica anterior se conocen varias enzimas DHFR adecuadas y, por consiguiente, genes que se pueden utilizar junto con la presente invención. La DHFR puede ser una DHFR natural o una variante funcional o derivado de esta. El término una «variante» o «derivado» incluye enzimas DHFR que tienen uno o más cambios en su secuencia de aminoácidos (p. ej., deleciones, sustituciones o adiciones) con respecto a la secuencia de aminoácidos de la enzima DHFR respectiva, proteínas de fusión que comprenden una enzima DHFR o fragmento funcional de esta y enzimas DHFR que han sido modificadas para proporcionar una estructura y/o función adicional, así como fragmentos funcionales de las anteriores, que siguen teniendo al menos una función de una enzima DHFR. El gen DHFR se selecciona preferentemente del grupo que consiste en DHFR natural, una variante de DHFR que tiene una menor sensibilidad a MTX en comparación con DHFR natural y una variante de DHFR que tiene una mayor sensibilidad a MTX en comparación con DHFR natural. De acuerdo con una realización, el gen DHFR es el gen DHFR natural. De acuerdo con una realización diferente, el gen DHFR codifica una variante menos funcional de DHFR. Esto conduce a una mayor sensibilidad a MTX y condiciones de selección más restrictivas. Esto se puede conseguir, p. ej., ubicando un intrón en el extremo 3' del gen DHFR (véase anteriormente). Una alternativa diferente podría basarse en el uso de un mutante/variante de DHFR que tiene una mayor sensibilidad a MTX que la DHFR natural. Estas realizaciones son particularmente beneficiosas si se utiliza el vector en células hospedadoras que son DHFR<sup>-</sup>.

Si se desea utilizar el sistema DHFR en células hospedadoras que incorporan una copia del gen DHFR en su propio genoma (p. ej., CHO DHFR<sup>+</sup>), se prefiere que se utilice un mutante/variante del gen DHFR, que es menos sensible a antifolatos tales como MTX que la DHFR natural y, por tanto, es en cierta medida resistente a antifolatos respectivamente MTX. P. ej., debido a mutaciones en el gen, la sensibilidad del gen DHFR a MTX se puede reducir significativamente de modo que dicha variante se puede utilizar en concentraciones de antifolato (MTX) más altas. Dicha variante de DHFR menos sensible a antifolatos y en particular MTX puede, p. ej., poseer una afinidad de unión a antifolato/MTX reducida. Se pueden utilizar variantes de DHFR respectivas para «sobrealorar» la expresión de DHFR endógena en líneas celulares naturales. Las respectivas variantes de DHFR «resistentes» o menos sensibles son muy conocidas en la técnica anterior.

El vector de expresión se puede seleccionar del grupo que consiste en

(a) un ácido nucleico de vector lineal que comprende los siguientes elementos genéticos en la disposición indicada, donde la dirección 5' a 3' se indica mediante la →:

- I. Promotor del casete de expresión (MASM) (→)
- II. Gen que codifica el marcador seleccionable amplificable de mamífero del casete de expresión (MASM) (→)
- III. Opcionalmente un intrón del casete de expresión (MASM) (→)
- IV. Sitio de poliA del casete de expresión (MASM) (→)
- V. Promotor del casete de expresión (POI) (→)
- VI. Intrón del casete de expresión (POI) (→)
- VII. Polinucleótido que codifica un polipéptido de interés, que se inserta en el casete de expresión (POI) (→)
- VIII. Sitio de poliA del casete de expresión (POI) (→)
- IX. Promotor del casete de expresión (POI') (→)
- X. Intrón del casete de expresión (POI') (→)
- XI. Polinucleótido que codifica un polipéptido de interés adicional, que se inserta en el casete de expresión (POI') (→)
- XII. Sitio de poliA del casete de expresión (POI') (→)
- XIII. Promotor del casete de expresión (MSM) (→)
- XIV. Gen que codifica el marcador seleccionable de mamífero del casete de expresión (MSM) (→)
- XV. Sitio de poliA del casete de expresión (MSM) (→)

## XVI. Casete de expresión PSM (→) o (←)

(b) un ácido nucleico de vector como el que se muestra como la Seq. ID No. 1 o Seq. ID No. 16 o un derivado de estas, que comprende la misma configuración respectivamente disposición de elementos genéticos.

La variante (a) en su forma circular que comprende un sitio de restricción para la linealización se muestra en la Fig. 1 y la correspondiente Tabla 1.

El vector linealizado de acuerdo con la presente invención se puede obtener disponiendo los casetes de expresión en el orden y la orientación apropiados tal como se describe en detalle anteriormente. La disposición de los casetes de expresión/elementos genéticos se puede realizar utilizando enzimas de restricción adecuadas y estrategias de clonación con el fin de ensamblar el vector de expresión. También se da a conocer un método para producir un ácido nucleico de vector como los que se dan a conocer en la presente, donde dicho método comprende disponer al menos los siguientes elementos genéticos

(a) al menos un casete de expresión (POI) para expresar un polipéptido de interés;

(b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero;

(c) un casete de expresión (MASM) que comprende un gen marcador seleccionable, amplificable de mamífero;

de tal modo que el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está situado en 3' respecto al casete de expresión (POI), y donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3'.

El vector de acuerdo con la presente invención se puede utilizar para expresar el polipéptido de interés en muchas células hospedadoras de mamífero diferentes. El vector de expresión linealizado de la presente invención normalmente se integra en el genoma y se mantiene en él. Hay dos formatos principales de células hospedadoras, cultivos de células adherentes y cultivos en suspensión. Se prefieren los cultivos en suspensión. Las líneas celulares más establecidas mantienen su carácter dependiente del anclaje a menos que se tomen medidas especiales para adaptarlas al crecimiento en suspensión. Las formulaciones de medios comercialmente disponibles facilitan la transición. Básicamente se pueden utilizar cualesquiera células hospedadoras de mamífero junto con la presente invención siempre que permitan la expresión de un polipéptido. Las células hospedadoras de mamífero adecuadas para los efectos de la presente invención incluyen, pero no se limitan a, células derivadas de ratones (p. ej., COP, L, C127, Sp2/0, NS-0, NS-1, At20 o NIH3T3), ratas (PC12, PC12h, GH3, MtT), hámsteres (p. Ej., BHK, CHO y CHO defectuosas para el gen DHFR), monos (p. ej., COS1, COS3, COS7, CV1 y Vero) y seres humanos (p. ej., Hela, HEK-293, PER-C6 derivadas de retina, células derivadas de fibroblastos diploides, células de mieloma y HepG2). Preferentemente, la célula hospedadora es una célula CHO. La construcción del vector de acuerdo con la presente invención es particularmente adecuada para producir polipéptidos en células de roedores tales como células CHO y CHO defectuosas para el gen DHFR.

También se proporcionan células hospedadoras de mamífero, que comprenden el vector de expresión linealizado de acuerdo con la presente invención. También se proporciona una línea celular estable, que comprende un vector de expresión linealizado de acuerdo con la presente invención o un segmento de este en el genoma. El segmento debe comprender al menos los casetes de expresión decisivos para la presente invención. Como el vector y sus características, así como las células hospedadoras adecuadas, se describen en detalle anteriormente, nos referimos a la divulgación anterior. Por consiguiente, también se proporciona un método para producir una célula hospedadora de mamífero tal como se ha descrito anteriormente, donde la célula hospedadora de mamífero se transfecta de forma estable con el ácido nucleico de vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 8.

Existen varios métodos apropiados conocidos en la técnica anterior para introducir un vector de expresión en una célula hospedadora de mamífero. Los métodos respectivos incluyen, pero no se limitan a, la transfección mediante fosfato de calcio, electroporación, lipofección, transferencia de genes mediada por polímeros y biolística. Las células hospedadoras adecuadas se han descrito anteriormente.

Después de introducir el ácido nucleico del vector de expresión en la célula o células hospedadoras, los transformantes obtenidos se cultivan en condiciones selectivas adecuadas para evaluar la expresión del gen marcador seleccionable de mamífero contenido en el casete de expresión (MSM). Esto significa que, por ejemplo, cuando el gen marcador seleccionable de mamífero es un gen de resistencia a antibiótico, se cultivan los transformantes en un medio que contiene el correspondiente antibiótico activo en células de mamífero y se seleccionan los transformantes que son viables en tales condiciones, lo cual posibilita de este modo la obtención de transformantes que expresan el gen marcador y, por tanto, han incorporado el vector. Además, se puede llevar a cabo un segundo paso de selección cultivando los transformantes en un medio de selección adaptado para seleccionar el gen marcador seleccionable, amplificable comprendido en el

casete de expresión (MASM). P. ej., si se utiliza DHFR como gen marcador seleccionable, amplificable, los transformantes se pueden cultivar en un medio exento de purina o nucleótidos en presencia de un inhibidor de DHFR.

5 Si se utiliza un promotor inducible en al menos un casete de expresión, se debería proporcionar una señal de inducción correspondiente con el fin de iniciar la expresión del polipéptido.

10 Con el fin de hacer uso del sistema de selección/amplificación de DHFR, dichas células hospedadoras se pueden cultivar en presencia de un inhibidor de DHFR. Los inhibidores de DHFR adecuados son antifolatos tales como, p. ej., MTX. La concentración de antifolato/MTX utilizada depende de la célula hospedadora y la variante de DHFR incorporada en el vector. El intervalo de concentración se puede elegir para los procedimientos de amplificación en múltiples pasos, en células hospedadoras DHFR<sup>-</sup>, por ejemplo, en valores de aproximadamente 5 nM - 20 nM que alcanzan valores de 500 nM a 1000 nM o incluso superiores para pasos de amplificación secundarios o adicionales. Para las células DHFR<sup>+</sup>, las concentraciones iniciales son habitualmente mayores en el intervalo de 100 nM a 750 nM, preferentemente 500 nM en los primeros pasos y de 500 nM a 1000 nM, y superiores para los pasos de amplificación adicionales. Se han descrito 15 variantes de DHFR adecuadas anteriormente.

20 Con el fin de aprovechar el sistema de selección/amplificación de GS, dichas células hospedadoras se pueden cultivar en presencia de, p. ej., MSX. La concentración de MSX utilizada depende de la célula hospedadora. El intervalo de concentración se puede elegir entre de aproximadamente 15 a 150 µM, de 20 a 100 µM y de 25 a 50 µM. Estos intervalos son particularmente adecuados para las células NSO y CHO.

25 Con el vector de expresión de acuerdo con la presente invención, se pueden expresar/producir varios polipéptidos de interés diferentes. El término polipéptido se refiere a una molécula que comprende un polímero de aminoácidos unidos entre sí por uno o más enlaces peptídicos. Los polipéptidos incluyen polipéptidos de cualquier longitud, que incluyen proteínas (p. ej., que tienen más de 50 aminoácidos) y péptidos (p. ej., de 2-49 aminoácidos). Los polipéptidos incluyen proteínas y/o péptidos de cualquier actividad o bioactividad, que incluyen, p. ej., polipéptidos bioactivos tales como proteínas o péptidos enzimáticos (p. ej., proteasas, cinasas, fosfatasa), proteínas o péptidos receptores, proteínas o péptidos transportadores, proteínas de unión a endotoxina y/o bactericidas, proteínas o péptidos estructurales, polipéptidos inmunitarios, toxinas, antibióticos, hormonas, factores de crecimiento, vacunas o similares. Dicho polipéptido 30 se puede seleccionar del grupo que consiste en hormonas peptídicas, interleucinas, activadores tisulares del plasminógeno, citocinas, inmunoglobulinas, en particular anticuerpos o fragmentos de anticuerpo o variantes de estos. Dicha inmunoglobulina puede ser de cualquier isotipo. Es muy común que se produzcan/necesiten moléculas de IgG (p. ej., IgG1) como proteínas terapéuticas. Un fragmento de anticuerpo es cualquier fragmento de un anticuerpo que comprenda al menos 20 aminoácidos de dicho anticuerpo entero, preferentemente al menos 100 aminoácidos, que al menos siga teniendo capacidad de unión a antígeno. El fragmento de anticuerpo puede comprender la región de unión 35 del anticuerpo tal como un fragmento Fab, un fragmento F(ab)<sub>2</sub>, multicuerpos que comprenden múltiples dominios de unión tales como diacuerpos, triacuerpos o tetracuerpos, aficuerpos o anticuerpos monodominio. Una variante de anticuerpo es un derivado de un anticuerpo o fragmento de anticuerpo que tiene la misma función de unión pero, p. ej., una secuencia de aminoácidos alterada. Dicho anticuerpo y/o fragmento de anticuerpo puede comprender una cadena ligera murina, cadena ligera humana, cadena ligera humanizada, cadena pesada humana y/o cadena pesada murina así como fragmentos activos o derivados de estas. Por ende, puede ser, p. ej., murino, humano, quimérico o humanizado. 40

45 La presente invención también proporciona métodos de producción de un polipéptido de interés, comprendiendo dicho método cultivar al menos una célula hospedadora de mamífero de acuerdo con la reivindicación 9 o 10 o una célula hospedadora de mamífero obtenida de acuerdo con la reivindicación 11 o 12 en un medio de cultivo celular en condiciones que permitan la expresión de dicho polipéptido de interés.

50 En un paso siguiente, dicho polipéptido se puede aislar/extraer a partir del cultivo celular. El polipéptido de interés expresado se puede obtener sometiéndolo a las células hospedadoras a disrupción. Los polipéptidos también se pueden expresar, p. ej., secretar al medio de cultivo y se pueden obtener a partir de este. También son posibles combinaciones de los respectivos métodos. De este modo, se pueden producir productos, en particular polipéptidos, y obtener/aislar de forma eficiente con un rendimiento alto. El polipéptido obtenido también se puede someter a pasos de procesamiento adicionales tales como, p. ej., pasos de purificación y/o modificación con el fin de producir el producto de interés con la calidad deseada. 55

60 De acuerdo con una alternativa, dicho polipéptido de interés se secreta al medio de cultivo celular y posteriormente se aísla a partir del medio de cultivo celular. El polipéptido es preferentemente una molécula de inmunoglobulina tal como un anticuerpo o un fragmento funcional de este. Con el fin de promover la secreción del polipéptido de interés se puede utilizar una secuencia líder. Preferentemente, se utiliza la secuencia líder de una molécula de inmunoglobulina.

El vector de expresión de acuerdo con la presente invención, así como las células hospedadoras de mamífero y polipéptidos de interés adecuados se describen en detalle anteriormente; nos referimos a la divulgación anterior.

El método para producir el polipéptido de interés puede comprender al menos uno de los siguientes pasos:

- aislar el polipéptido de interés a partir de dicho medio de cultivo celular y/o a partir de dicha célula hospedadora; y/o
- procesar el polipéptido de interés aislado.

El polipéptido de interés producido de acuerdo con la invención se puede recuperar, después purificar, aislar y/o modificar mediante métodos conocidos en la técnica. Por ejemplo, el producto se puede recuperar a partir del medio nutritivo mediante procedimientos convencionales que incluyen, pero no se limitan a, centrifugación, filtración, ultrafiltración, extracción o precipitación. La purificación se puede llevar a cabo mediante diversidad de procedimientos conocidos en la técnica que incluyen, pero no se limitan a, cromatografía (p. ej., de intercambio iónico, afinidad, hidrofóbica, cromatoenfoco y exclusión por tamaño), procedimientos electroforéticos (p. ej., enfoque isoelectrico preparativo), solubilidad diferencial (p. ej., precipitación con sulfato de amonio) o extracción.

De acuerdo con una realización que es en particular ventajosa para la producción de proteínas/péptidos farmacéuticos, la célula hospedadora se cultiva en suspensión en condiciones sin suero. El polipéptido obtenido se puede purificar posteriormente, p. ej., purificando el polipéptido presente en el sobrenadante del cultivo celular utilizando métodos cromatográficos (p. ej., purificación por afinidad).

Los polipéptidos producidos de acuerdo con el método de la presente invención presentan buenas propiedades de estabilidad. Los resultados también muestran que los polipéptidos se expresan de una forma funcional y, por ende, en la conformación correcta. Por consiguiente, también se dan a conocer polipéptidos obtenidos mediante el método de producción de acuerdo con la presente invención utilizando el vector de expresión descrito en detalle anteriormente. Tal como se ha expuesto anteriormente, los polipéptidos se obtienen con un buen rendimiento. El polipéptido es preferentemente una molécula de inmunoglobulina tal como un anticuerpo o un fragmento funcional de este.

La **Fig. 1** muestra un ácido nucleico de vector circular.

La **Fig. 2** muestra la versión linealizada del ácido nucleico de vector de acuerdo con la Fig. 1 con el fin de demostrar la influencia de la posición del sitio de restricción para la linealización.

Los números 1 a 17 mostrados en las Figs. 1 y 2 indican los elementos/características genéticas del ácido nucleico de vector y se describen en detalle en la tabla 1 subsiguiente. Si no se definen de otro modo, las flechas blancas caracterizan promotores o elementos promotores/potenciadores; los recuadros rayados caracterizan elementos intrónicos, las flechas negras simbolizan los polinucleótidos para expresar el polipéptido de interés; los recuadros cuadriculados caracterizan el sitio de poliA; las flechas cuadriculadas los genes marcadores de mamífero del casete de expresión (MSM) y (MASM); la flecha rayada el gen marcador seleccionable procarionótico. Como se puede observar, todos los elementos genéticos están dispuestos en la misma orientación 5' a 3' (indicada por la dirección de la flecha). Los ácidos nucleicos de vector pBW147, pBW154 y pBW160, que se describen en más detalle a continuación, se construyen respectivamente.

**Tabla 1 - orientación y disposición de los elementos genéticos de acuerdo con un ácido nucleico de vector tal como se muestra en las Fig. 1 y 2**

Numeración en las Fig. 1 y 2	Elemento genético
1	Promotor del casete de expresión (POI) (representado por la flecha blanca). Es, p. ej., un promotor/potenciador de CMV.
2	Intrón del casete de expresión (POI) (representado por el recuadro rayado). Es, p. ej., un intrón RK, tal como se ha descrito anteriormente.
3	Polinucleótido que codifica un polipéptido de interés, que se inserta en el casete de expresión (POI) (representado por la flecha negra). De acuerdo con la realización mostrada, es la cadena ligera de un anticuerpo monoclonal (mAB-LC).
4	Sitio de PoliA del casete de expresión (POI) (representado por el recuadro cuadriculado). Es, p. ej., un sitio de PoliA de SV40.

5	Promotor del casete de expresión (POI') (representado por la flecha blanca). Es, p. ej., un promotor/potenciador de CMV.
6	Intrón del casete de expresión (POI') (representado por el recuadro rayado). Es, p. ej., un intrón RK, tal como se ha descrito anteriormente.
7	Polinucleótido que codifica un polipéptido de interés adicional, que se inserta en el casete de expresión (POI') (representado por la flecha negra). De acuerdo con la realización mostrada, es la cadena pesada de un anticuerpo monoclonal (mAB-HC).
8	Sitio de PoliA del casete de expresión (POI') (representado por el recuadro cuadrículado). Es, p. ej., un sitio de PoliA de SV40.
9	Promotor del casete de expresión (MSM) (representado por la flecha blanca). Es, p. ej., un promotor/potenciador de SV40.
10	Gen que codifica el marcador seleccionable de mamífero del casete de expresión (MSM) (representado por la flecha cuadrículada). Es, p. ej., el gen neo.
11	Sitio de PoliA del casete de expresión (MSM) (representado por la barra). Es, p. ej., un sitio de PoliA sintético tal como se ha descrito anteriormente.
12	Casete de expresión PSM (representado por la flecha rayada). Puede, p. ej., comprender un gen marcador seleccionable procariótico que proporcione una resistencia frente a la ampicilina.
13	Sitio de restricción para la linealización (representado por la barra). Dicho sitio es preferentemente un sitio de restricción único.
14	Promotor del casete de expresión (MASM) (representado por la flecha blanca). Es, p. ej., un promotor de SV40.
15	Gen que codifica el marcador seleccionable amplificable de mamífero del casete de expresión (MASM) (representado por la flecha cuadrículada). Es, p. ej., el gen DHFR.
16	Intrón en el casete de expresión (MASM) (representado por la barra). Este intrón está presente opcionalmente.
17	Sitio de PoliA del casete de expresión (MASM) (representado por el recuadro cuadrículado). Es, p. ej., un sitio de PoliA de SV40.

Subsiguientemente, se proporcionan ejemplos adecuados para los elementos del vector descritos, que, sin embargo, no son limitantes.

- 5 Como marcador seleccionable amplificable de mamífero se prefiere el uso de DHFR. Se proporciona un ejemplo adecuado de un polinucleótido de DHFR de ratón natural con la Seq. ID No. 5 que se utiliza preferentemente junto con las células hospedadoras DHFR<sup>-</sup>. Se proporciona una forma mutada adecuada de DHFR con la Seq. ID No. 6. Una forma mutada respectiva se utiliza preferentemente junto con las células DHFR<sup>+</sup>. La Seq. ID No. 12 muestra un mutante de DHFR que incluye un intrón adecuado para incrementar aún más la presión de selección (véase anteriormente). También se pueden utilizar variantes o fragmentos funcionales de las anteriores.

Como gen marcador seleccionable de mamífero se prefiere el uso de *neo*. Se proporciona una secuencia adecuada con la Seq. ID No. 7. También se pueden utilizar variantes o fragmentos funcionales de esta.

- 15 Como secuencia promotora para dirigir la expresión del polipéptido de interés, se prefiere el uso de un promotor de CMV. Se proporciona una secuencia adecuada con la Seq. ID No. 8. También se pueden utilizar variantes o fragmentos funcionales de esta.

- 20 Como secuencia promotora para dirigir la expresión de los genes marcadores seleccionables MSM y MASM, se prefiere el uso de un promotor de SV40. Se proporciona una secuencia adecuada con la Seq. ID No. 9. También se pueden utilizar variantes o fragmentos funcionales de esta.

- 25 Como secuencia de poliadenilación para los polipéptidos de interés y/o el MASM se puede utilizar un sitio de poliA de SV40. Se muestra una secuencia adecuada como la Seq. ID No. 10. También se pueden utilizar variantes o fragmentos funcionales de esta o una orientación invertida (sitio de poliA tardío o temprano de SV40).

- Como secuencia intrónica para el casete de expresión (POI) que codifica el polipéptido de interés, se puede utilizar un intrón Rk. Se proporciona una secuencia adecuada con la Seq. ID No. 11. También se pueden utilizar variantes o fragmentos funcionales de esta.

- 30 Un sitio de poliadenilación sintético que se puede utilizar, p. ej., junto con el marcador seleccionable de mamífero (MSM) se muestra como la Seq. ID No. 13. También se pueden utilizar variantes o fragmentos funcionales de esta.

Un marcador seleccionable bacteriano (PSM) adecuado es, p. ej., el gen de la beta-lactamasa que proporciona resistencia a la ampicilina. Se proporciona una secuencia adecuada con la Seq. ID No. 14. También se pueden utilizar variantes o fragmentos funcionales de esta.

Además, el ácido nucleico de vector puede comprender al menos un sitio de clonación múltiple (MCS) para insertar, p. ej., un polinucleótido que codifica un polipéptido de interés. MCS se puede proporcionar en 3' y 5' del polinucleótido que codifica el polipéptido de interés. Se proporcionan sitios MCS adecuados como la Seq. ID No. 4 (preferentemente situada en el sitio/región 5') y la Seq. ID No. 15 (preferentemente situada en el sitio/región 3'). Estos sitios MCS se pueden utilizar, p. ej., con el fin de introducir el polinucleótido que codifica el polipéptido de interés.

Se muestran ácidos nucleicos de vector particularmente preferidos como la Seq. ID No. 1 (que comprende el gen DHFR natural, dicho vector es particularmente útil para el sistema DHFR-) y la Seq. ID No. 16 (que comprende un gen DHFR mutado, dicho vector es particularmente útil para el sistema DHFR+).

## EJEMPLOS

A continuación, se describe la presente invención por medio de ejemplos no limitantes, que constituyen, sin embargo, realizaciones preferidas de la presente invención. Los ejemplos que no conciernen a la invención reivindicada son únicamente a efectos ilustrativos.

### I. Métodos de cultivo celular y transfección

Subsiguientemente, se describen, por medio de ejemplos, métodos apropiados para transfectar y cultivar las células hospedadoras de acuerdo con la presente invención para expresar un polipéptido de interés.

#### Ejemplo 1: Cultivo celular

Se cultivan células CHO en un medio de CHO adecuado tal como, p. ej., ExCell81134 (obtenido en SAFC Biosciences). Las células se pasan 2-3 veces por semana a medio fresco y se mantienen en fase de crecimiento logarítmico a lo largo del estudio.

#### Ejemplo 2: Estrategia de transfección

Para la transfección, se utilizan células CHO progenitoras en fase de crecimiento exponencial con una viabilidad superior al 90%. Las transfecciones por lipofección se realizan utilizando el reactante DMRIE-C de acuerdo con las instrucciones del fabricante (Invitrogen). La cantidad de células se ajusta a  $1 \times 10^5$  células en medio OptiMEM I (Invitrogen). Para la lipofección, se mezclan 2  $\mu\text{g}$  o 4  $\mu\text{g}$  del vector de expresión y 4  $\mu\text{l}$  del reactivo DMRIE-C durante 28 min a temperatura ambiente y se añaden a las células durante 4 h a 37°C. Las células se diluyen entonces hasta  $2 \times 10^5$  células/ml de medio de cultivo y se incuban durante 2 días a 37°C y un 5 % de CO<sub>2</sub>.

#### Ejemplo 3: Selección con neomicina y amplificación génica

El marcador de selección de neomicina situado en el ácido nucleico del vector de expresión permite seleccionar la resistencia a G418. Para seleccionar los transfectantes, las células se cultivan en presencia de 0,8 mg/ml de G418 (Invitrogen) durante aproximadamente dos semanas. Dos semanas después de la transfección y selección con G418, emergen poblaciones agrupadas que consisten predominantemente en células resistentes a G418. A continuación se cultivan las células en ausencia de nucleótidos durante aproximadamente dos semanas. Entonces se inicia la amplificación génica mediante la adición de MTX 20 nM al medio de cultivo. Después de tres semanas de cultivo, se genera una población de células heterogéneas amplificadas. El marcador de selección/amplificación de DHFR (dihidrofolato reductasa) permite la amplificación del gen DHFR así como del transgén mediante la adición del análogo al ácido fólico metotrexato (MTX) a los medios de cultivo, lo cual da como resultado títulos mayores para las poblaciones de la transfección. Después de recuperarse de la crisis, las poblaciones se someten entonces a un segundo y tercer paso de amplificación utilizando una concentración mayor de MTX, cada uno durante aproximadamente dos semanas (MTX 100 nM y 500 nM). En cada paso, las células se congelan después de la recuperación de las poblaciones.

#### Ejemplo 4: Establecimiento de líneas celulares clonales

Para obtener una línea celular clonal (es decir, una línea celular derivada de una única célula), la población de células transfectadas de forma estable se puede diluir y sembrar en placas de 96 pocillos con una densidad celular de 0,3 - 0,5 células por pocillo (dilución limitada). Se aumenta la escala de las células que forman una colonia distinta utilizando procedimientos estándar. Eventualmente, se evalúan los clones individuales para determinar la expresión del polipéptido recombinante, siendo los mayores productores retenidos después del cultivo y análisis. Entre estos candidatos, se elige

la línea celular con las características de crecimiento y productividad apropiadas para la producción de la proteína recombinante. La productividad habitualmente se puede mejorar aún más estableciendo/adaptando las condiciones de cultivo, es decir, añadiendo aditivos tales como peptonas.

5 **II. Construcciones de los vectores**

10 Son factibles diversos ensamblajes de vectores de acuerdo con las enseñanzas de la presente invención. Como los elementos individuales del vector son conocidos en la técnica anterior, se pueden ensamblar vectores adecuados, p. ej., mediante secuenciación o amplificación y la clonación apropiada de los elementos genéticos básicos y casetes de expresión en la orientación deseada. Los respectivos métodos de clonación son estado de la técnica y en la técnica anterior también se describe la secuencia de los elementos genéticos descritos anteriormente. Subsiguiente, se describe la generación de varios constructos de vectores a modo de ejemplo. Sin embargo, los expertos en la técnica entienden que otras realizaciones y formas diferentes de obtener los respectivos vectores son adecuadas y que se puede disponer de ellas fácilmente.

15 Con el fin de facilitar la comprensión de la disposición de los constructos de vectores descritos en la presente a modo de ejemplo y sus vectores precursores, las tablas 1 y 2 proporcionan un resumen sobre los elementos genéticos principales comprendidos en estos vectores, su orden y orientación. Evidentemente, sólo se muestran los elementos principales, los vectores, sin embargo, pueden comprender secuencias de esqueleto o elementos genéticos adicionales. Cada columna en la tabla representa un constructo de vector. Desde la fila de arriba hasta la fila de abajo, se enumeran los elementos genéticos de los casetes de expresión en el orden de su disposición y su orientación en el vector. Como los vectores descritos son circulares, el elemento que se muestra en la última fila de cada columna es, de hecho, adyacente al elemento genético que se muestra en la primera fila (evidentemente, puede haber secuencias de esqueleto presentes). La orientación de cada elemento genético está indicada por las flechas. La punta de la flecha apunta a la dirección 3' del elemento genético respectivo.

20 **Tabla 2 - orden de los elementos genéticos en los vectores precursores**

pBW133 10416bp	pBW139 9213bp	pBW146 9247bp	pBW159 11122bp
Prom/pot de CMV →	Prom/pot de CMV →	Prom/pot de CMV, que comprende un sitio de restricción de <i>SpeI</i> (153) →	Prom/pot de CMV →
Intrón RK →	Intrón RK →	Intrón RK →	Intrón RK →
mAB – LC →	mAB – LC →	mAB-LC →	mAB-LC →
MCS I	MCS	MCS	MCS
PoliA de SV40→	PoliA de SV40→	PoliA de SV40→	PoliA de SV40→
Sitio de restricción de <i>DraIII</i> (2365)	Prom/pot de CMV →	Prom/pot de CMV que comprende un sitio de restricción de <i>SpeI</i> (2519) →	Prom/pot de CMV →
DHFR* que porta un sitio de restricción de <i>Scal</i> (2870); orientación inversa ←	Intrón RK →	Intrón RK →	Intrón RK →
Prom/pot de SV40; orientación inversa ←	MCS	MCS2	mAB-HC →
Sitio de restricción de <i>DraIII</i> (3651)	mAB-HC →	mAB-HC →	MCS
Prom/pot de SV40 →	MCS	MCS	Promotor de T3
Neo →	PoliA de SV40→	Promotor de T3	PoliA de SV40→
Prom/pot de CMV →	Región del fago f1 que comprende un sitio de restricción de <i>DraIII</i> (5456) →	PoliA de SV40→	Región del fago f1 →
Intrón RK →	Prom/pot de SV40 →	Región del fago f1 →	Prom/pot de SV40 →
MCS2	Neo →	Prom/pot de SV40 que comprende un origen	Neo →

		mínimo de replicación de SV40 →	
mAB-HC →	PoliA de SV40→	Neo →	PoliA sint
PoliA de SV40→	Amp que comprende un sitio de restricción de <i>ScaI</i> (7828) →	PoliA sint	Amp →
Amp que porta <i>ScaI</i> (9031) →	Sitio de restricción de <i>BglII</i> (9209), 5' respecto al prom/pot de CMV	Amp →	3 sitios de pA
		Sitio de restricción de <i>BglII</i> (9243), 5' adyacente al prom/pot de CMV	intrón
			DHFR; orientación inversa ←
			Prom de SV40; orientación inversa ←
			Sitio de restricción de <i>SwaI</i> (11113)

**Tabla 3 - orden de los elementos genéticos en los vectores de expresión**

<b>pBW147</b> <b>11053bp</b>	<b>pBW154</b> <b>11109bp</b>	<b>pBW160</b> <b>11122bp</b>
Prom/pot de CMV, que comprende un sitio de restricción de <i>SpeI</i> (153) →	Prom/pot de CMV →	Prom/pot de CMV →
Intrón RK →	Intrón RK →	Intrón RK →
mAB-LC →	mAB-LC →	mAB-LC →
PoliA de SV40→	PoliA de SV40→	PoliA de SV40→
Prom/pot de CMV que comprende un sitio de restricción de <i>SpeI</i> (2519) →	Prom/pot de CMV →	Prom/pot de CMV →
Intrón RK →	Intrón RK →	Intrón RK →
mAB-HC →	mAB-HC →	mAB-HC →
PoliA de SV40→	PoliA de SV40→	PoliA de SV40→
Prom/pot de SV40 →	Prom/pot de SV40 →	Prom/pot de SV40 →
Neo →	Neo →	Neo →
PoliA sint	PoliA sint	PoliA sint
Amp →	Amp →	Amp →
Sitio de restricción de <i>SwaI</i> (9288)	Sitio de restricción de <i>SwaI</i> (9243)	Sitio de restricción de <i>SwaI</i> (9256)
Prom/pot de SV40 →	Prom de SV40 →	Prom de SV40 →
DHFR* →	DHFR →	DHFR →
Sitio de pA de Bgh →	Intrón	Intrón
	PoliA de SV40	PoliA de SV40

5 Las abreviaturas en las tablas 1 a 3 anteriores y en las Figs. 1 y 2 tienen los significados habituales como será evidente para un experto en la técnica y tal como se han descrito anteriormente, y tienen en particular los siguientes significados:

MCS = sitio de clonación múltiple

10 mAB-HC = cadena pesada de anticuerpo monoclonal

mAB-LC = cadena ligera de anticuerpo monoclonal

intrón = véase Grillari *et al.*, 2001, 3. *Biotechnol.* 87, 59-65

prom/pot = promotor/potenciador

5 Ejemplo 5: Construcción del vector de expresión pBW147

10 En este procedimiento, se evalúa la configuración en tándem de genes de mAB y DHFR\* (variante mutada que tiene una sensibilidad menor a MTX que la DHFR natural) combinada con el sitio de pA de bgh. El casete de DHFR\* está ubicado en 5' antes del casete de expresión (POI), que comprende la mAB-LC, de modo que todos los marcos abiertos de lectura estén ubicados en una dirección de lectura. El ensamblaje de pBW147 se muestra en la tabla 3.

pBW147 se puede construir a partir de pBW133 (remítase a la tabla 2). La construcción de pBW147 se describe en la presente.

15 *Construcción de pBW133*

Las construcciones de vectores descritas a continuación se basan en el vector de expresión comercializado pCI-neo (Promega Cooperation, EE. UU.). La secuencia de ADN completa es de acceso público (Número de acceso en el GenBank/EMBL: U47120). Se introduce un sitio de clonación múltiple nuevo en pCIneo.

20 Las dos hebras del sitio de clonación múltiple se sintetizan *de novo*. pCIneo se corta con *NheI* y *XmaI*. El MCS viejo se elimina mediante electroforesis en gel. El sitio de clonación múltiple nuevo se sintetiza de modo que los 4 nucleótidos terminales del extremo 5' de la hebra de antisentido y el extremo 3' terminal de la hebra de ADN superior no se sinteticen. Después de la hibridación de ambas hebras, se crean extremos compatibles para *NheI* y *XmaI*.

25 La secuencia de este sitio de clonación múltiple nuevo es la siguiente (véase la Seq. ID No. 2):

<i>ApalSgrAI</i>				
<i>Agel</i>	<i>PmeI</i>	<i>EcoRV</i>	<i>PshAI</i>	
<i>EcoO109I</i>	<i>BstEII</i>	<i>PmlI</i>	<i>BspEI</i>	<i>Ascl</i>

30 El plásmido resultante de la unión de pCIneo con el MCS nuevo se considera pCI-neo-2 a efectos descriptivos. pCI-neo-2 se modifica adicionalmente mediante la introducción del intrón pRK de pRK5 (BD PharMingen). Por lo tanto, pCI-neo-2 se digiere con *Apal*. Se crean extremos romos mediante el tratamiento con polimerasa T4. A continuación, el plásmido se digiere con *NdeI*. pRK5 se digiere con *NdeI* y *NruI* (cortadora de extremos romos). El fragmento que contiene el intrón RK se aísla y se une con el esqueleto de pCIneo2. El plásmido resultante es pCIneo2RK.

35 Para juntar ambos casetes de expresión en un vector, se puede obtener el vector pCIneoDHFR\*-RK. pCIneoDHFR\*-RK se obtiene como sigue:

40 El casete de expresión de DHFR\* se amplifica mediante PCR a partir del vector pCHI-LC (vector de expresión de cadena ligera de SP2/0 Simulect). Los cebadores son BB35 (GGGCACTACGTGCCGCGGATTTAAATGCGGCCGCATATGGTGCCT – Seq. ID No. 3) y BB36 (GGGCACGTAGTGTATTAGGGGAGCAGAGAACTTGAA – Seq. ID No. 4).

45 El fragmento de PCR se clona en pCIneoRK a través de la digestión con *DraIII* de restricción para dar el vector pCIneoDHFR\*-RK. pCIneoDHFR\*-RK se abre por digestión con *EcoO109I*. A fin de crear extremos romos, se realiza posteriormente un tratamiento con la enzima Klenow.

50 El casete de expresión de pCIneo2RK se extirpa digiriendo el plásmido con *BglII*, *NgoMIV* y *StuI*. Después de unir los dos fragmentos, se crea el vector de expresión «vacío» pCHO2neoN. En pCHO2neoN se inserta el gen de la cadena ligera de anticuerpo a través del sitio de restricción de *MluI* y *SaII* creando de este modo un constructo de vector, al que nos referimos como pBW108. El gen de anticuerpo se amplifica, por lo tanto, utilizando cebadores que contienen los dos sitios de restricción.

55 La cadena pesada de mAB se inserta en pBW108 a través de la digestión con *PmeI* y *Ascl* del vector, para obtener de este modo un constructo de vector que se considera pBW111 a efectos descriptivos. La cadena pesada se amplifica por PCR con lo cual el extremo 5' genera extremos romos y el extremo 3' del gen contiene el sitio de restricción de *Ascl*.

En pBW111, el extremo de la región no traducida 5' de la cadena ligera se reemplaza porque hay un codón ATG adicional presente delante del ADNc de la cadena ligera. Esto se consigue extirpando un fragmento *BglII* / *MluI* de pBW111 y

reemplazándolo con el fragmento correcto. El nuevo plásmido se considera pBW133. pBW133 es el primer vector con todos los genes en un plásmido. La disposición de los genes es: LC – DHFR (dirección opuesta) – neo – HC (véase también la tabla 2). Este vector es uno de los materiales de partida que se pueden utilizar para obtener vectores de acuerdo con las enseñanzas de la presente invención. Sin embargo, es obvio que existen otras formas diversas de obtener vectores respectivos.

#### Construcción de pBW139

El segundo constructo de vector que se puede utilizar para obtener pBW147 tiene la configuración de pBW139. pBW139 se puede crear a partir de pBW115. Para la construcción de pBW115, se clona el gen de la cadena pesada en pCIneoRK (véase anteriormente). Por lo tanto, pCIneoRK se digiere con *MluI* y *NruI* (cortadora de extremos romos), mientras que el fragmento de PCR de la cadena pesada se digiere con *AscI* (3') (compatible con *MluI*) y es romo en 5'. El plásmido resultante se considera pBW115.

pBW115 se digiere con *Scal* y *BglII*. A continuación, se realiza un relleno con Klenow para crear extremos romos. El casete de expresión de la cadena ligera se extirpa de pBW133 con *Scal* y *DraIII*. A fin de crear extremos romos, se realiza un tratamiento con ADN·polimerasa T4. Como resultado de la unión se obtiene un vector que tiene una configuración como la de pBW139 (véase la tabla 2).

#### Ensamblaje de pBW147

Para obtener pBW147, se digiere pBW133 con *SpeI*, *XhoI*. El fragmento que contiene partes del promotor de CMV y la primera parte de la cadena pesada se aísla y se une a pBW139, que también se corta con *SpeI*, *XhoI*. En el vector resultante, el casete de la cadena pesada se encuentra sin un codón ATG adicional perturbador. A fin de devolver la cadena ligera al vector, se digiere pBW139 con *SpeI*. El fragmento que contiene la LC se inserta en pBW148 que se abre con *SpeI*. El plásmido resultante tiene una configuración como la de pBW146 (véase la tabla 2).

En pBW146, se inserta el gen DHFR de pBW112 (vector de expresión de otro proyecto). Sin embargo, el gen DHFR también se podría haber obtenido de una fuente diferente, dependiendo del tipo de variante de DHFR deseada. pBW146 se digiere con *BglII* y *BamHI*. Posteriormente, el casete de DHFR se amplifica por PCR con cebadores que contienen los sitios de restricción de *BglII* y *BamHI*. El fragmento de PCR se digiere con las dos enzimas y se inserta en el sitio de *BglII* apropiado de pBW146. El plásmido resultante tiene una configuración como la de pBW147, donde todos los casetes de expresión tienen la misma orientación. La estructura se muestra en la tabla 3.

Este vector de expresión se puede utilizar para obtener transfecciones estables. A fin de aumentar aún más el rendimiento de la expresión, se puede elegir una variante de DHFR natural muy «sensible» a MTX, en la que también se deben de adaptar las concentraciones de MTX de forma adecuada.

#### Ejemplo 6: Construcción del vector de expresión pBW154

Para este vector, se evalúan la configuración en tándem de los genes de mAB y el casete del gen DHFR de pSV2DHFR (versión natural de DHFR con alta sensibilidad a MTX). El casete de expresión de DHFR de pSV2DHFR (ATCC#374146) se amplifica por PCR. El fragmento contenía el promotor y los sitios de poliA. Al igual que antes, los oligos tenían sitios de restricción de *BglII* / *BamHI*. El casete de expresión de DHFR se inserta en el sitio de restricción de *BglII* de pBW146, lo cual da como resultado un constructo de vector que tiene la misma estructura que pBW154. La estructura de pBW154 se muestra en la tabla 3 y se puede derivar de las Fig. 1 y 2 que muestran un ejemplo general de un constructo de vector que tiene una estructura/configuración general respectiva de los elementos genéticos. La secuencia de pBW154 se proporciona como la Seq. ID No 1. El polinucleótido de la cadena ligera se marca como n (indicado en la solicitud de prioridad con el marcador de posición V), el polinucleótido de la cadena pesada se marca como n (indicado en la solicitud de prioridad con el marcador de posición Y) en la Seq. ID No. 1. Las características de pBW154 se resumen en la Tabla 4. Evidentemente, también se pueden utilizar otros elementos del vector, p. ej., promotores diferentes, potenciadores diferentes, sitios de PoliA diferentes y otros elementos tales como oris diferentes. Además, es posible intercambiar los casetes de expresión para la cadena ligera y la pesada de la molécula de inmunoglobulina. Sin embargo, se prefiere la selección y disposición mostrada de los elementos del vector. Tal como se indica anteriormente, también se pueden utilizar fragmentos funcionales de moléculas de inmunoglobulina. Por lo tanto, las «nnn» indicadas solamente tienen carácter ilustrativo y no indican ninguna restricción de tamaño ya que puede haber secuencias de inmunoglobulina más grandes o más pequeñas presentes en la posición correspondiente. Con el fin de facilitar la comparación con las Figs. 1 y 2, que muestran la construcción general de vectores de la presente divulgación, hemos indicado la numeración de los elementos correspondientes en las Figs. 1 y 2.

#### **Tabla 4**

Inicio de los pares de bases	Fin	Característica	Numeración correspondiente en las Fig. 1 y 2
1	743	Prom/pot de CMV	1
857	1000	Intrón RK	2
1054	1766	mAB LC	3
1815	2036	PoliA de SV40	4
2367	3109	Prom/pot de CMV	5
3223	3366	Intrón RK	6
3452	4863	mAB HC	7
4931	5152	PoliA de SV40	8
5766	6184	Prom de SV40	9
6229	7024	Neomicina fosfotransferasa	10
7087	7135	PoliA sintético	11
7546	8406	Gen de resistencia a antibiótico betalactámico	12
9243		Sitio de linealización único	13
9422	9617	Prom de SV40	14
9640	10204	DHFR	15
9776	10426	Intrón (Donante – Aceptor)	16
10909	11098	PoliA de SV40	17

5 Todos los elementos genéticos están dispuestos en la misma orientación 5' a 3'. Con este constructo de vector que utiliza una DHFR natural, la amplificación génica tal como se ha descrito anteriormente funciona de forma muy eficiente. El título en un experimento por lotes estándar se puede incrementar en 10-20 veces. En la presente, se observa un gran incremento en el título de expresión de anticuerpos tras la amplificación con MTX. Partiendo del tratamiento con G418, a lo largo del tratamiento sin nucleótidos hasta diversas concentraciones diferentes con MTX (MTX 20 y 100 nM), el título aumenta de forma constante y considerable. Sin embargo, la utilización de concentraciones mucho más altas de MTX (p. ej., MTX 500 nM) no suele aportar ventajas adicionales con células CHO, aunque se pueden utilizar concentraciones altas respectivas. Los títulos de anticuerpos obtenidos para poblaciones a lo largo del proceso de selección/amplificación variaban de 2 a más de 60 mg/L cuando se utilizaban procedimientos de cultivo estándar. Los títulos de expresión se pueden incrementar aún más al establecer líneas celulares clonales a partir de las poblaciones y utilizar medios personalizados para mejorar la expresión celular ya que el título obtenible también depende el medio utilizado.

#### 15 Ejemplo 7: Construcción del vector de expresión pBW160

15 Los experimentos también demuestran que la orientación del gen DHFR en el vector es decisiva. En pBW146 (véase anteriormente), hay un sitio de restricción de *EcoRI* presente. Con el fin de que *EcoRI* sea la única cortadora presente en el vector de expresión final, el sitio se puede eliminar dirigiendo pBW146 con *EcoRI*, relleno de Klenow y relegación. Esto da como resultado un plásmido que tiene una configuración como la de pBW158 (no se muestra). El casete de DHFR se puede integrar en pBW158, tal como se ha descrito anteriormente. Dado que ambas orientaciones (orientación como la que se muestra en pBW159 y pBW160, véanse las tablas 2 y 3) se generan de forma automática, se pueden evaluar ambas para determinar los niveles de expresión. Nuestros resultados muestran un rendimiento superior de la configuración con todos los marcos abiertos de lectura orientados en una dirección de lectura 5' a 3'.

25 Los vectores que tienen una configuración como la del vector pBW159 (véase la tabla 2), donde la orientación de DHFR es en orden inverso de los genes de mAB habitualmente muestran solamente títulos de expresión muy bajos, incluso después de la amplificación con MTX (habitualmente menos de 1mg/L). Los vectores que tienen un diseño como el de pBW160 (véase la tabla 3), donde la orientación de DHFR está en el marco con los genes de mAB, pueden proporcionar títulos de anticuerpos más altos de más de 5 mg/L e incluso más de 10 o incluso más de 15 mg/L (obtenibles a partir de las poblaciones). Nuevamente, al establecer una línea celular clonal y al utilizar un medio de alto rendimiento, el rendimiento de título se puede incrementar aún más cuando se utilizan los constructos de vectores de acuerdo con la presente invención.

35 Estos experimentos pueden demostrar la ventaja de la «configuración en el marco» de los marcadores de selección y los genes que codifican mAB tal como se utilizan de acuerdo con las enseñanzas de la presente invención. Este resultado respalda el descubrimiento de que la orientación 5' a 3' de los elementos del vector es un factor importante para vectores de alta expresión. Además, la estabilidad de la expresión es muy favorable con los vectores de expresión de acuerdo con la presente invención.

La configuración del vector de acuerdo con la presente invención permite la generación sencilla de poblaciones celulares con productividades altas específicas para las células. Son elementos clave la orientación 5' a 3', las variantes de DHFR elegidas y la ubicación del marcador de selección de DHFR en el vector así como la disposición de los genes de anticuerpo y el segundo marcador de selección (neo) en el plásmido. El vector también se puede utilizar para la producción de proteínas o péptidos que no sean anticuerpos. Tal como se ha descrito anteriormente, con ligeras adaptaciones del casete de DHFR, este sistema también se puede utilizar para la amplificación génica en la línea celular CHO-K1PD positiva para DHFR. Para la amplificación génica en células hospedadoras DHFR<sup>+</sup>, se utiliza una versión mutada del gen DHFR (véase anteriormente). El casete de expresión de DHFR completo de un vector que comprende una versión mutada del gen DHFR tal como pBW117 se puede amplificar por PCR con cebadores que incorporen un sitio de *Bam*HI. Este fragmento se clona entonces en el sitio de *Bgl*II de pBW158 lo cual da como resultado el vector pBW165. Con vectores que tienen una configuración como la de pBW165 que comprenden un gen DHFR mutado (y continúa con otros anticuerpos), se pueden generar líneas celulares de alto rendimiento en la célula hospedadora CHO-K1-PD, que es una línea celular DHFR<sup>+</sup>. Un ejemplo de una secuencia de vector respectiva se proporciona como la Seq. ID No. 16. Evidentemente, también se pueden utilizar elementos del vector diferentes a los que se muestran, p. ej., promotores diferentes, potenciadores diferentes, sitios de PolIA diferentes y/u otros elementos tales como, p. ej., oris diferentes. Además, es posible intercambiar los casetes de expresión para la cadena ligera y la pesada de la molécula de inmunoglobulina. Sin embargo, se prefiere la selección y disposición mostrada de los elementos del vector. Como se indica anteriormente, a partir del vector se pueden expresar moléculas de inmunoglobulina completas así como fragmentos funcionales de moléculas de inmunoglobulina. En la Seq. ID No. 16 solamente se indica como sitio de inserción el sitio donde se puede situar/insertar la secuencia de inmunoglobulina respectiva en el vector de expresión final. Cualesquiera secuencias de inmunoglobulina pueden estar presentes en la posición correspondiente. Es más, como se indica anteriormente, también es posible expresar polipéptidos de interés diferentes.

Ejemplo 8: Producción a pequeña escala de anticuerpos con células CHO transfectadas

Para el análisis en cultivos en suspensión, se siembran células con una densidad de 1x10<sup>5</sup> células/ml en 50 ml de medio ExCell81134 (SAFC Biosciences) en un matraz de cultivo de fondo redondo de 250 ml con tapa de filtro. Las células se agitan a 65 rpm en un agitador/incubador ISF-4-W de Kühner a 37°C en un entorno de un 10 % de CO<sub>2</sub> a lo largo de la duración del estudio. Se proporcionan alimentaciones individuales con soluciones patentadas de acuerdo con un esquema de alimentación fijo que comienza el 4.º día de la expansión celular. El 13.º día, se extraen muestras de 1 ml y se mide el título utilizando HPLC estándar y una columna de proteína A.

El sobrenadante del cultivo celular resultante de los cultivos en matraces oscilantes de los mejores clones se purifica mediante cromatografía de afinidad con proteína A.

Ejemplo 9: Purificación con proteína A del anticuerpo expresado

Para la purificación con proteína A se cargan aproximadamente 27 mL de sobrenadante de cultivo sin células que contienen aproximadamente 32,4 mg de anticuerpo en una columna de afinidad MabSelect de 0,5x10 cm. Tras la carga, la columna se enjuaga y lava suficientemente. A continuación se eluye el anticuerpo a pH 3-4. El eluato se analiza con HPLC estándar utilizando una columna de proteína A. Se obtienen aproximadamente 30,5 mg de anticuerpo.

**III. Ejemplos de productividades específicas para las células y rendimientos**

Los clones que se seleccionan después de la expansión clonal se evalúan para determinar su productividad.

Ejemplo 10: Expresión de un anticuerpo IgG1

Se expresa un anticuerpo IgG1. Se cultivan los clones en el medio comercializado ExCell81134 (SAFC Biosciences). Se añaden soluciones de alimentación y aditivos convencionales tales como peptona. Se pueden obtener tasas de productividad altas cuando se utiliza el vector de acuerdo con al presente invención:

Clon	Qp (pg/célula/día)
1	114
2	91
3	103

Qp = productividad específica para las células.

Ejemplo 11: Expresión de un anticuerpo IgG1 y un anticuerpo IgG4

Se expresan un anticuerpo IgG1 y un anticuerpo IgG4. Los clones se cultivan en un medio de cultivo apropiado. Se añaden soluciones de alimentación y aditivos convencionales tales como peptona. Se pueden obtener tasas de productividad altas cuando se utiliza el vector de acuerdo con al presente invención:

5

Clon	Anticuerpo IgG1 Qp (pg/célula/día)	Anticuerpo IgG4 Qp (pg/célula/día)
1	76	
2	96	
3		73

**Qp** = productividad específica para las células.

10 Ejemplo 12: Producto a gran escala de polipéptidos con células CHO transfectadas

La producción de polipéptidos a gran escala se puede llevar a cabo, por ejemplo, en biorreactores de agitación por ondas, de vidrio o de acero inoxidable. Con ese fin, las células se expanden, habitualmente partiendo de un único vial congelado, por ejemplo, un vial de un banco de células patrón. Las células se descongelan y se expanden a lo largo de varios pasos. Se inoculan biorreactores de diferente escala con cantidades apropiadas de células. La densidad celular se puede aumentar añadiendo soluciones de alimentación y aditivos al biorreactor. Las células se mantienen con una viabilidad alta durante un periodo prolongado. A gran escala se consiguen títulos de producto en el reactor que oscilan de unos pocos cientos de miligramos por litro a varios gramos por litro. La purificación se puede llevar a cabo mediante metodología de cromatografía estándar, que puede incluir pasos de cromatografía de afinidad, intercambio iónico, interacción hidrófoba o exclusión por tamaño. El tamaño del biorreactor puede ser de un volumen de hasta varios miles de litros en la escala final (véase también, p. ej., F. Wurm, *Nature Biotechnology* Vol. 22, 11, 2004, 1393-1398).

15

20

Ejemplo 13: Estrategia de clonación para introducir nuevos genes de anticuerpo en los vectores

25

Una estrategia - entre otras - para insertar nuevos polipéptidos de interés es la siguiente (explicada a modo de ejemplo utilizando un vector que tiene una configuración como la de pBW154):

*Clonación del gen de la cadena ligera*

30

El gen de la cadena ligera se puede amplificar por PCR con cebadores que introducen un sitio de *MluI* 5' del codón ATG y un sitio de *SalI* 3' del gen. El producto de PCR se introduce en pBW154 a través de estas dos enzimas de restricción. Esto conduce a un vector intermedio que consiste en solamente la cadena ligera.

*Clonación del gen de la cadena pesada*

35

El gen de la cadena pesada se puede amplificar por PCR con cebadores que introducen un extremo romo para la utilización del sitio de *NruI* 5' del codón ATG del vector y un sitio de *XbaI* 3' del gen. El producto de PCR se introduce en pBW154 a través de estas dos enzimas de restricción. Esto conduce a un vector intermedio con la cadena ligera vieja y la cadena pesada nueva.

40

*Ensamblaje del vector final*

El fragmento que contiene la mAB-HC nueva se extirpa del vector de HC a través de la digestión con *SalI*. Se inserta entonces en el vector intermedio de LC a través de *SalI* para dar como resultado el nuevo vector de LC-HC final.

45

## LISTA DE SECUENCIAS

<110> Novartis AG  
 <120> Vector de expresión de mamífero  
 <130> 57 019 K  
 <150> EP07150339.5  
 <151> 2007-12-21  
 <160> 16  
 <170> PatentIn versión 3.5  
 <210> 1  
 <211> 11109  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Secuencia del vector pBW154  
 <220>  
 <221> característica miscelánea  
 <222> (1)..(743)  
 <223> Prom/pot de CMV  
 <220>  
 <221> característica miscelánea  
 <222> (825)..(1024)  
 <223> Intrón RK  
 <220>  
 <221> característica miscelánea  
 <222> (857)..(1000)  
 <223> Región del intrón RK que se elimina tras la expresión  
 <220>  
 <221> característica miscelánea  
 <222> (1054)..(1767)  
 <223> mAB LC (indicada originalmente en la Seq. Id. No: 1 con el símbolo «v»)  
 <220>  
 <221> característica miscelánea  
 <222> (1815)..(2036)  
 <223> SV40polyA  
 <220>  
 <221> característica miscelánea  
 <222> (2367)..(3109)  
 <223> Prom/pot de CMV  
 <220>  
 <221> característica miscelánea  
 <222> (3191)..(3390)  
 <223> Intrón RK  
 <220>  
 <221> característica miscelánea  
 <222> (3223)..(3366)  
 <223> Intrón RK que se elimina tras la expresión  
 <220>  
 <221> característica miscelánea  
 <222> (3452)..(4864)  
 <223> mAB HC (indicada originalmente en la Seq. Id. No: 1 con el símbolo «y»)  
 <220>  
 <221> característica miscelánea  
 <222> (4931)..(5152)

<223> PoliA de SV40  
 <220>  
 <221> característica miscelánea  
 <222> (5247)..(5702)  
 <223> Región f1  
 <220>  
 <221> característica miscelánea  
 <222> (5766)..(6184)  
 <223> Prom de SV40  
 <220>  
 <221> característica miscelánea  
 <222> (6229)..(7023)  
 <223> Gen de la neomicina·fosfotransferasa  
 <220>  
 <221> característica miscelánea  
 <222> (7087)..(7135)  
 <223> PoliA sintético  
 <220>  
 <221> característica miscelánea  
 <222> (7546)..(8406)  
 <223> Gen de resistencia a antibiótico betalactámico  
 <220>  
 <221> característica miscelánea  
 <222> (9243)..(9243)  
 <223> Sitio de linealización único  
 <220>  
 <221> característica miscelánea  
 <222> (9276)..(9623)  
 <223> SV40prom  
 <220>  
 <221> característica miscelánea  
 <222> (9521)..(9568)  
 <223> Origen mínimo de replicación de SV40  
 <220>  
 <221> característica miscelánea  
 <222> (9640)..(10203)  
 <223> Gen de DHFR  
 <220>  
 <221> característica miscelánea  
 <222> (9776)..(10426)  
 <223> Intrón (Donador - Aceptor)  
 <220>  
 <221> característica miscelánea  
 <222> (10910)..(11097)  
 <223> SV40polyA  
 <400> 1

ES 2 784 926 T3

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
ttggccattg catacgttgt atctatatca taatatgtac atttatatg gctcatgtcc	120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg	180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
gcctggctga ccgcccaacg acccccgcc attgacgtca ataatgacgt atgtcccat	300
agtaacgccca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc	360
ccacttgga gtacatcaag tgtatcatat gccaaagtacg cccctattg acgtcaatga	420
cggtaaatgg cccgcctggc attatgccca gtacatgacc ttatgggact ttcctactg	480
gcagtacatc tacgtattag tcatcgctat taccatgggtg atgcggtttt gccagtacat	540
caatgggctg ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt	600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc	660
cgccccattg acgcaaatgg gcggtaggcg tgtacgggtg gaggtctata taagcagagc	720
tcgttttagtg aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag	780
aagacaccgg gaccgatcca gcctccgagg ccgggaacgg tgcattggaa cgcggattcc	840
ccgtgccaa agtgacgtaa gtaccgccta tagagtctat aggccacccc ccttggttc	900
gtagaacgc ggctacaatt aatacataac cttatgtatc atacacatac gatttaggtg	960
acactataga ataacatcca ctttgccttt ctctccacag gtgtccactc ccaggtccaa	1020
ctgcacctcg gttctatcga aaacgcgtcc accnnnnnnn nnnnnnnnnn nnnnnnnnnn	1080
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1140
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1200
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1260
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1320
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1380
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1440
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1500
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1560
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1620
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1680
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn	1740
nnnnnnnnnn nnnnnnnnnn nnnnnnnngtc gaccgggagc gccgctccc tttagtgagg	1800

ES 2 784 926 T3

gttaatgctt cgagcagaca tgataagata cattgatgag tttggacaaa ccacaactag 1860  
 aatgcagtga aaaaaatgct ttatttgtga aatttgtgat gctattgctt tatttghtaac 1920  
 cattataagc tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt 1980  
 tcagggggag atgtgggagg ttttttaaag caagtaaac ctctacaaat gtggtaaaat 2040  
 ccgataagga tcgatccggg ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc 2100  
 caacagttgc gcagcctgaa tggcgaatgg acgcgccctg tagcggcgca ttaagcgcgg 2160  
 cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcctca gcgcccgtc 2220  
 ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa 2280  
 atcgggggct ccctttaggg ttccgattta gagctttacg gcacctcgac cgcaaaaaac 2340  
 ttgatttggg tgatggttca cgatcttcaa tattggccat tagccatatt attcattggt 2400  
 tatatagcat aaatcaatat tggctattgg ccattgcata cgttgatctc atatcataat 2460  
 atgtacattt atattgctc atgtccaata tgaccgccat gttggcattg attattgact 2520  
 agttattaat agtaatcaat tacgggggtca ttagttcata gcccatatat ggagttccgc 2580  
 gttacataac ttacggtaaa tggcccgctt ggctgaccgc ccaacgacc ccgcccattg 2640  
 acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa 2700  
 tgggtggagt atttacggt aactgccac ttggcagtac atcaagtga tcatatgcca 2760  
 agtaccccc ctattgacgt caatgacggg aatggcccg cctggcatta tgcccagtac 2820  
 atgaccttat gggactttcc tacttggcag tacatctacg tattagtcac cgctattacc 2880  
 atgggtgatgc ggttttggca gtacatcaat gggcgtggat agcggtttga ctcacgggga 2940  
 tttccaagtc tccaccccat tgacgtcaat gggagtttgt tttggacca aaatcaacgg 3000  
 gactttccaa aatgtcgtaa caactccgcc ccattgacgc aaatgggagg taggcgtgta 3060  
 cgggtggagg tctatataag cagagctcgt ttagtgaacc gtcagatcgc ctggagacgc 3120  
 catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct ccgcgccgg 3180  
 gaacggtgca ttggaacgcg gattccccgt gccaaagatg acgtaagtac cgcctataga 3240  
 gtctataggc ccacccctt ggcttcgtta gaacgcggct acaattaata cataacctta 3300  
 tgtatcatac acatacgatt taggtgacac tatagaataa catccacttt gcctttctct 3360  
 ccacaggtgt ccaactccag gtccaactgc acctcggttc tatcgcgatt gaattccccg 3420  
 gggatcctct aggggtgaccg tttgtgccac cnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3480  
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3540  
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3600  
 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 3660

ES 2 784 926 T3

nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4260
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4320
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4680
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4740
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	4860
nnnnnggcgcg	tggtacctct	agagtcgacc	cgggcgcccg	cttccttta	gtgagggtta	4920
atgcttcgag	cagacatgat	aagatacatt	gatgagttt	gacaaaccac	aactagaatg	4980
cagtgaaaaa	aatgctttat	ttgtgaaatt	tgtgatgcta	ttgctttatt	tgtaaccatt	5040
ataagctgca	ataaacaagt	taacaacaac	aattgcattc	atthttatg	tcaggttcag	5100
ggggagatgt	gggaggtttt	ttaaagcaag	taaaacctct	acaaatgtgg	taaaatccga	5160
taaggatcga	tccgggctgg	cgtaatagcg	aagaggcccg	caccgatcgc	ccttccaac	5220
agttgcgcag	cctgaatggc	gaatggacgc	gccctgtagc	ggcgcattaa	gcgcggcggg	5280
tgtggtggtt	acgcgcagcg	tgaccgctac	acttgcagc	gccctagcgc	ccgctccttt	5340
cgctttcttc	ccttcctttc	tcgccacggt	cgccggcttt	ccccgtcaag	ctctaaatcg	5400
ggggctccct	ttagggttcc	gatttagagc	tttacggcac	ctcgaccgca	aaaaacttga	5460
tttgggtgat	ggttcacgta	gtgggccatc	gccctgatag	acgggttttc	gccctttgac	5520
gttggagtcc	acgttcttta	atagtggact	cttgttccaa	actggaacaa	cactcaaccc	5580

ES 2 784 926 T3

tatctcggtc tattcttttg atttataagg gattttgccg atttcggcct attggttaa	5640
aaatgagctg atttaacaaa tatttaacgc gaattttaac aaaatattaa cgtttacaat	5700
ttcgcctgat gcggtatfff ctctttacgc atctgtgcgg tatttcacac cgcatacgcg	5760
gatctgcgca gcaccatggc ctgaaataac ctctgaaaga ggaacttggg taggtacctt	5820
ctgaggcgga aagaaccagc tgtggaatgt gtgtcagtta ggggtgtgaa agtccccagg	5880
ctccccagca ggcagaagta tgcaaagcat gcatctcaat tagtcagcaa ccagggtgtg	5940
aaagtcccca ggctccccag caggcagaag tatgcaaagc atgcatctca attagtcagc	6000
aaccatagtc cgcgccctaa ctccgcccat cccgcccta actccgccca gttccgccca	6060
ttctccgccc catggctgac taatfttttt tatttatgca gaggccgagg ccgcctcggc	6120
ctctgagcta ttccagaagt agtgaggagg cttttttgga ggcctaggct tttgcaaaaa	6180
gcttgattct tctgacacaa cagtctcga cttaggcta gagccaccat gattgaacaa	6240
gatggattgc acgcaggttc tccggccgct tgggtggaga ggctattcgg ctatgactgg	6300
gcacaacaga caatcggctg ctctgatgcc gccgtgttcc ggctgtcagc gcaggggagc	6360
ccggttcttt ttgtcaagac cgacctgtcc ggtgcctga atgaactgca ggacgaggca	6420
gcgcggctat cgtggctggc cacgacgggc gttccttgcg cagctgtgct cgacgttctc	6480
actgaagcgg gaagggactg gctgctattg ggcgaagtgc cggggcagga tctcctgtca	6540
tctcaccttg ctctcgccga gaaagtatcc atcatggctg atgcaatgcg gcggctgcat	6600
acgcttgatc cggctacctg cccattcagc caccaagcga aacatcgcac cgagcgagca	6660
cgtactcgga tgggaagccg tcttgtcgat caggatgatc tggacgaaga gcatcagggg	6720
ctcgcgccag ccgaactggt cgcacggctc aaggcgcgca tgcccagcgg cgaggatctc	6780
gtcgtgacct atggcgatgc ctgcttgccg aatatcatgg tggaaaatgg ccgcttttct	6840
ggattcatcg actgtggccg gctgggtgtg gcggaccgct atcaggacat agcgttggct	6900
accogtgata ttgctgaaga gcttggcggc gaatgggctg accgcttcct cgtgctttac	6960
ggtatcgccg ctcccgatc gcagcgcac gccttctatc gccttcttga cgagttcttc	7020
tgagcgggac tctggggttc gaaatgaccg accaagcgcg gcccaacctg ccatcacgat	7080
ggccgcaata aaatatcttt attttcatta catctgtgtg ttggtttttt gtgtgaatcg	7140
atagcgataa ggatccgctg atggtgcact ctcagtacaa tctgctctga tgccgcatag	7200
ttaagccagc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgctc	7260
ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatgtg tcagaggttt	7320
tcaccgctcat caccgaaacg cgcgagacga aagggcctcg tgatacgcct atttttatag	7380
gttaatgtca tgataataat ggtttcttag acgtcagggtg gcacttttctg gggaaatgtg	7440

ES 2 784 926 T3

cgcggaaccc ctatttgttt atttttctaa atacattcaa atatgtatcc gctcatgaga 7500  
 caataaccct gataaatgct tcaataatat tgaaaaagga agagtatgag tattcaacat 7560  
 ttccgtgtcg cccttattcc cttttttgcg gcattttgcc ttcctgtttt tgctcaccca 7620  
 gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc 7680  
 gaactggatc tcaacagcgg taagatcctt gagagttttc gccccgaaga acgttttcca 7740  
 atgatgagca cttttaagat tctgctatgt ggcgcggtat tatcccgtat tgacgccggg 7800  
 caagagcaac tcggtgcgcc catacactat tctcagaatg acttggttga gtactcacca 7860  
 gtcacagaaa agcatcttac ggatggcatg acagtaagag aattatgcag tgctgccata 7920  
 accatgagtg ataacactgc ggccaactta cttctgacaa cgatcggagg accgaaggag 7980  
 ctaaccgctt ttttgacaaa catgggggat catgtaactc gccttgatcg ttgggaaccg 8040  
 gagctgaatg aagccatacc aaacgacgag cgtgacacca cgatgcctgt agcaatggca 8100  
 acaacgttgc gcaaactatt aactggcgaa ctacttactc tagcttcccg gcaacaatta 8160  
 atagactgga tggaggcggg taaagttgca ggaccacttc tgcgctcggc ccttccggct 8220  
 ggctggttta ttgctgataa atctggagcc ggtgagcgtg ggtctcggg tatcattgca 8280  
 gcactggggc cagatggtaa gccctcccgt atcgtagtta tctacacgac ggggagtcag 8340  
 gcaactatgg atgaacgaaa tagacagatc gctgagatag gtgcctcact gattaagcat 8400  
 tgtaactgt cagaccaagt ttactcatat atactttaga ttgatttaaa acttcatttt 8460  
 taatttaaaa ggatctaggt gaagatcctt tttgataatc tcatgaccaa aatcccttaa 8520  
 cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga 8580  
 gatccttttt ttctgcgctt aatctgctgc ttgcaaacaa aaaaaccacc gctaccagcg 8640  
 gtggtttgtt tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc 8700  
 agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccaactcaag 8760  
 aactctgtag caccgcctac atacctcgtt ctgctaatac tgttaccagt ggctgctgcc 8820  
 agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg 8880  
 cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttgagcgg aacgacctac 8940  
 accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttc cgaagggaga 9000  
 aagcgggaca ggtatccggg aagcggcagg gtcggaacag gagagcgcac gagggagctt 9060  
 ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag 9120  
 cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg 9180  
 gcctttttac ggttccctggc cttttgctgg ccttttgctc acatggctcg acagatccat 9240  
 ttaaattttc accgtcatca ccgaaacgcg cgaggcagct gtggaatgtg tgtcagttag 9300  
 ggtgtggaaa gtccccaggc tccccagcag gcagaagtat gcaaagcatg catctcaatt 9360

ES 2 784 926 T3

agtcagcaac caggtgtgga aagtccccag gctccccagc aggcagaagt atgcaaagca 9420  
 tgcatctcaa ttagtcagca accatagtcg cccccctaac tccgcccatac cggcccctaa 9480  
 ctccgcccag ttccgcccata tctccgcccc atggctgact aatttttttt atttatgcag 9540  
 aggccgagggc cgcctcgcc tctgagctat tccagaagta gtgaggaggc ttttttgag 9600  
 gcctaggctt ttgcaaaaag ctttatcccc gctgccatca tggttcgacc attgaactgc 9660  
 atcgtgccc tgtcccaaga tatggggatt ggcaagaacg gagacctacc ctggcctccg 9720  
 ctgaggaacg agttcaagta cttccaaaga atgaccacaa cctcttcagt ggaaggtaaa 9780  
 cagaatctgg tgattatggg taggaaaacc tggttctcca ttcctgagaa gaatcgacct 9840  
 ttaaaggaca gaattaatat agttctcagt agagaactca aagaaccacc acgaggagct 9900  
 cattttcttg ccaaaagttt ggatgatgcc ttaagactta ttgaacaacc ggaattggca 9960  
 agtaaagtag acatggtttg gatagtcgga ggcagttctg tttaccagga agccatgaat 10020  
 caaccaggcc acctcagact ctttgtagaca aggatcatgc aggaatttga aagtgcacag 10080  
 tttttccag aaattgattt ggggaaatat aaacttctcc cagaataacc aggcgtcctc 10140  
 tctgaggtcc aggaggaaaa aggcataag tataagtttg aagtctacga gaagaaagac 10200  
 taacaggaag atgctttcaa gttctctgct cccctcctaa agctatgcat tttataaga 10260  
 ccatgggact tttgctggct ttagatcttt gtgaaggaac cttacttctg tgggtgtaga 10320  
 taattggaca aactacctac agagatttaa agctctaagg taaatataaa atttttaagt 10380  
 gtataatgtg taaactact gattctaatt gtttgtgtat tttagattcc aacctatgga 10440  
 actgatgaat gggagcagtg gtggaatgcc tttaatgagg aaaacctgtt ttgctcagaa 10500  
 gaaatgccat ctagtatga tgaggctact gctgactctc aacattctac tcctccaaaa 10560  
 aagaagagaa aggtagaaga cccaaggac tttcctcag aattgctaag ttttttgagt 10620  
 catgctgtgt ttagtaatag aactcttgct tgctttgcta tttacaccac aaaggaaaa 10680  
 gctgcactgc tatacaagaa aattatggaa aatattctg taacctttat aagtaggcat 10740  
 aacagttata atcataacat actgtttttt cttactccac acaggcatag agtgtctgct 10800  
 attaataact atgctcaaaa attgtgtacc tttagctttt taatttgtaa aggggttaat 10860  
 aaggaatatt tgatgtatag tgccttgact agagatcata atcagccata ccacatttgt 10920  
 agaggtttta cttgctttaa aaaacctccc acacctccc ctgaacctga aacataaaat 10980  
 gaatgcaatt gttgtgtta acttgtttat tgcagcttat aatggttaca aataaagcaa 11040  
 tagcatcaca aatttcacaa ataaagcatt ttttctactg cattctagtt gtggtttgaa 11100  
 ttcgatct 11109  
 <210> 2  
 <211> 59  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Sitio I de MCS  
 <400> 2

ES 2 784 926 T3

ctagggccca ccggtgaccg tttaaacacg tgatatccgg acaattgtcg gcgcgccgg 59  
 <210> 3  
 <211> 46  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Cebador  
 <400> 3  
 gggcactacg tgccgcggat ttaaatgcgg ccgcatatgg tgcact 46  
 <210> 4  
 <211> 38  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Cebador  
 <400> 4  
 gggcacgtag tgtttattag gggagcagag aacttgaa 38  
 <210> 5  
 <211> 564  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> DHFR natural  
 <400> 5  
 atggttcgac cattgaactg catcgtcgcc gtgtoccaaag atatggggat tggcaagaac 60  
 ggagacctac cctggcctcc gctcaggaac gagttcaagt acttccaaag aatgaccaca 120  
 acctcttcag tggaaggtaa acagaatctg gtgattatgg gtaggaaaac ctggttctcc 180  
 attcctgaga agaatcgacc tttaaaggac agaattaata tagttctcag tagagaactc 240  
 aaagaaccac cacgaggagc tcattttctt gccaaaagtt tggatgatgc cttaaagactt 300  
 attgaacaac cggaattggc aagtaaagta gacatggttt ggatagtcgg aggcagttct 360  
 gtttaccagg aagccatgaa tcaaccaggc cacctcagac tctttgtgac aaggatcatg 420  
 caggaatttg aaagtgcac gtttttccca gaaattgatt tggggaaata taaacttctc 480  
 ccagaatacc caggcgtcct ctctgaggtc caggaggaaa aaggcatcaa gtataagttt 540  
 gaagtctacg agaagaaaga ctaa 564  
 <210> 6  
 <211> 564  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Mutante de DHFR  
 <400> 6

ES 2 784 926 T3

atggttcgac cattgaactg catcgtcgcc gtgtcccaa ataggggat tggcaagaac 60  
 ggagaccgac cctggcctcc gctcaggaac gagttcaagt acttccaaag aatgaccaca 120  
 acctcttcag tggaaggtaa acagaatctg gtgattatgg gtaggaaaac ctggttctcc 180  
 attcctgaga agaatcgacc tttaaaggac agaattaata tagttctcag tagagaactc 240  
 aaagaaccac cagcaggagc tcattttctt gccaaaagt tggatgatgc cttaagactt 300  
 attgaacaac cgggaattgac aagtaaagta gacatggttt ggatagtcgg aggcagttct 360  
 gtttaccagg aagccatgaa tcaaccaggc cacctcagac tctttgtgac aaggatcatg 420  
 caggaatttg aaagtgcac gtttttccca gaaattgatt tggggaaata taaacttctc 480  
 ccagaatacc caggcgtcct ctctgaggtc caggaggaaa aaggcatcaa gtataagttt 540  
 gaagtctacg agaagaaaga ctaa 564  
 <210> 7  
 <211> 795  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Gen de resistencia a neomicina  
 <400> 7  
 atgattgaac aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc 60  
 ggctatgact gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccgctgtca 120  
 gcgcaggggc gcccggttct ttttgcgaag accgacctgt ccggtgccct gaatgaactg 180  
 caggacgagg cagcgcggct atcgtggctg gccacgacgg gcgctccttg cgcagctgtg 240  
 ctcgacgttg tcaactgaagc gggaaaggac tggttgctat tgggcgaagt gccggggcag 300  
 gatctoctgt catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg 360  
 cggcggctgc atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc 420  
 atcgagcgag cacgtactcg gatggaagcc ggtcttgcg atcaggatga tctggacgaa 480  
 gagcatcagg ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgcac 540  
 ggcgaggatc tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat 600  
 ggccgctttt ctggattcat cgactgtggc cggtgggtg tggcggaccg ctatcaggac 660  
 atagcgttgg ctaccctgga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc 720  
 ctctgctttt acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt 780  
 gacgagttct tctga 795  
 <210> 8  
 <211> 743  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Promotor de CMV  
 <400> 8

ES 2 784 926 T3

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60  
 ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120  
 aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg 180  
 gtcattagtt catagcccat atatggagtt ccgcggttaca taacttacgg taaatggccc 240  
 gcctggctga ccgcccacg acccccgccc attgacgtca ataatgacgt atgttcccat 300  
 agtaacgccca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc 360  
 ccacttgcca gtacatcaag tgtatcatat gccaaagtacg ccccctattg acgtcaatga 420  
 cggtaaatgg cccgcctggc attatgcccga gtacatgacc ttatgggact ttcctacttg 480  
 gcagtacatc tacgtattag tcatcgctat taccatgggtg atgcggtttt ggcagtacat 540  
 caatgggctg ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt 600  
 caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaaatgtc gtaacaactc 660  
 cgccccattg acgcaaatgg gcggtaggcg tgtacgggtg gaggtctata taagcagagc 720  
 tcgttttagtg aaccgtcaga tcg 743  
 <210> 9  
 <211> 419  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Promotor de SV40  
 <400> 9  
 gcgcagcacc atggcctgaa ataacctctg aaagaggaac ttggtaggt accttctgag 60  
 gcggaaagaa ccagctgtgg aatgtgtgtc agttaggggtg tggaaagtcc ccaggctccc 120  
 cagcagggcag aagtatgcaa agcatgcac tcaattagtc agcaaccagg tgtggaaagt 180  
 ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 240  
 tagtcccgcc cctaactccg cccatcccgc ccctaactcc gcccagttcc gccattctc 300  
 cgccccatgg ctgactaatt tttttatatt atgcagaggc cgaggccgcc tcggcctctg 360  
 agctattcca gaagtgtga ggaggctttt ttggaggcct aggcttttgc aaaaagctt 419  
 <210> 10  
 <211> 222  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Sitio de poliA de SV40  
 <400> 10  
 cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 60  
 aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 120  
 ataaacaagt taacaacaac aattgcattc attttatggt tcagggtcag ggggagatgt 180  
 gggaggtttt ttaaagcaag taaaacctct acaaatgtgg ta 222  
 <210> 11  
 <211> 200  
 <212> ADN  
 <213> Artificial  
 <220>

ES 2 784 926 T3

<223> Intrón RK  
 <400> 11  
 ttggaacgcg gattccccgt gccaaagagt acgtaagtac cgcctataga gtctataggg 60  
 ccacccccctt ggcttcgcta gaacgcggct acaattaata cataacctta tgtatcatac 120  
 acatacgatt taggtgacac tatagaataa catccacttt gcctttctct ccacaggtgt 180  
 ccactcccag gtccaactgc 200  
 <210> 12  
 <211> 942  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Mutante de DHFR que incluye intrón  
 <220>  
 <221> característica miscelánea  
 <222> (565)..(942)  
 <223> intrón  
 <400> 12  
 atggttcgac cattgaactg catcgtcgcc gtgtccaaa atatggggat tggcaagaac 60  
 ggagaccgac cctggcctcc gctcaggaac gaggttcaagt acttccaaag aatgaccaca 120  
 acctcttcag tggaaggtaa acagaatctg gtgattatgg gtaggaaaac ctggttctcc 180  
 attcctgaga agaatcgacc tttaaaggac agaattaata tagttctcag tagagaactc 240  
 aaagaaccac cacgaggagc tcattttctt gccaaaagtt tggatgatgc cttaagactt 300  
 attgaacaac cggaattggc aagtaaagta gacatggttt ggatagtcgg aggcagttct 360  
 gtttaccagg aagccatgaa tcaaccaggc cacctcagac tctttgtgac aaggatcatg 420  
 caggaatttg aaagtgacac gtttttccca gaaattgatt tggggaaata taaacttctc 480  
 ccagaatacc caggcgtcct ctctgaggtc caggaggaaa aaggcatcaa gtataagttt 540  
 gaagtctacg agaagaaaga ctaacaggaa gatgctttca agttctctgc tccccctca 600  
 aagctatgca tttttataag accatggggg atgctcgatc ccctcgcgag ttggttcagc 660  
 tgctgcctga ggctggacga cctcgcggag ttctaccggc agtgcaaadc cgtcggcatc 720  
 caggaaacca gcagcggcta tccgcgcac catgccccg aactgcagga gtggggaggc 780  
 acgatggccg ctttgggtccg gatctttgtg aaggaacctt acttctgtgg tgtgacataa 840  
 ttggacaac tacctacaga gatttaaagc tctaaggtaa atataaaatt ttaagtgtg 900  
 taatgtgtta aactactgat tctaattggt tgtgtatttt ag 942  
 <210> 13  
 <211> 49  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Sitio de poliA sintético  
 <400> 13  
 aataaaatat ctttatcttc attacatctg tgtgttggtt ttttgtgtg 49  
 <210> 14  
 <211> 861  
 <212> ADN

ES 2 784 926 T3

<213> Artificial  
 <220>  
 <223> Gen de resistencia a antibiótico betalactámico  
 <400> 14  
 atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt ttgccttct 60  
 gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca gttgggtgca 120  
 cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag ttttcgcccc 180  
 gaagaacggt ttccaatgat gagcactttt aaagttctgc tatgtggcgc ggtattatcc 240  
 cgtattgacg ccgggcaaga gcaactcggg cgcgcatac actattctca gaatgacttg 300  
 gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt aagagaatta 360  
 tgcagtgctg ccataacatc gagtgataac actgcgcca acttacttct gacaacgatc 420  
 ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt aactcgcctt 480  
 gatcgttggg aaccggagct gaatgaagcc ataccaaacy acgagcgtga caccacgatg 540  
 cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact tactctagct 600  
 tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc acttctgcgc 660  
 tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct 720  
 cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac 780  
 acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc 840  
 tcaactgatta agcattggta a 861  
 <210> 15  
 <211> 39  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Sitio II de MCS  
 <400> 15  
 ctagcctcga gaattcacgc gtggtacctc tagagtcga 39  
 <210> 16  
 <211> 9210  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Secuencia del vector pBW165  
 <220>  
 <221> característica miscelánea  
 <222> (1)..(743)  
 <223> Prom/pot de CMV  
 <220>  
 <221> característica miscelánea  
 <222> (825)..(1024)  
 <223> Intrón RK  
 <220>  
 <221> característica miscelánea  
 <222> (857)..(1000)  
 <223> Región del intrón RK que se elimina tras la expresión  
 <220>  
 <221> variación

<222> (1053)..(1054)  
 <223> inserción de mAB LC o fragmento funcional de esta  
 <220>  
 <221> característica miscelánea  
 <222> (1101)..(1322)  
 <223> SV40polyA  
 <220>  
 <221> característica miscelánea  
 <222> (1653)..(2395)  
 <223> Prom/pot de CMV  
 <220>  
 <221> característica miscelánea  
 <222> (2477)..(2676)  
 <223> Intrón RK  
 <220>  
 <221> característica miscelánea  
 <222> (2509)..(2652)  
 <223> región del intrón RK que se empalma tras la expresión  
 <220>  
 <221> variación  
 <222> (2750)..(2751)  
 <223> inserción de mAB HC o fragmento funcional de esta  
 <220>  
 <221> característica miscelánea  
 <222> (2817)..(3038)  
 <223> SV40polyA  
 <220>  
 <221> característica miscelánea  
 <222> (3133)..(3588)  
 <223> Región f1  
 <220>  
 <221> característica miscelánea  
 <222> (3652)..(4070)  
 <223> Prom de SV40  
 <220>  
 <221> característica miscelánea  
 <222> (3968)..(4033)  
 <223> origen mínimo de replicación  
 <220>  
 <221> característica miscelánea  
 <222> (4115)..(4909)  
 <223> Gen de la neomicina·fosfotransferasa  
 <220>  
 <221> característica miscelánea  
 <222> (4973)..(5021)  
 <223> poliA sintético  
 <220>  
 <221> característica miscelánea  
 <222> (5432)..(6292)  
 <223> Resistencia a antibiótico betalactámico  
 <220>  
 <221> característica miscelánea  
 <222> (7170)..(7508)  
 <223> Prom de SV40  
 <220>

ES 2 784 926 T3

<221> característica miscelánea

<222> (7585)..(8148)

<223> Mutante de DHFR

<220>

<221> característica miscelánea

<222> (8149)..(8526)

<223> intrón de mutante de DHFR

<220>

<221> característica miscelánea

<222> (9010)..(9205)

<223> poliA de SV40

<400> 16

tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc	120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg	180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
gcctggctga ccgccaacg acccccggcc attgacgtca ataatgacgt atgttcccat	300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaacgtc	360
ccacttggca gtacatcaag tgtatcatat gccaaagtacg ccccctattg acgtcaatga	420
cggtaaatgg cccgcctggc attatgcca gtacatgacc ttatgggact ttcctacttg	480
gcagtacatc tacgtattag tcatcgtat taccatgggt atgcggtttt ggcagtacat	540
caatgggctg ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt	600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc	660
cgccccattg acgcaaatgg gcggtaggcg tgtacgggtg gaggtctata taagcagagc	720
tcgttttagtg aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag	780
aagacaccgg gaccgatcca gcctccgcgg ccgggaacgg tgcattggaa cgcggattcc	840
ccgtgccaaag agtgacgtaa gtaccgccta tagagtctat agggccacc ccttggcttc	900
gttagaacgc gggtacaatt aatacataac cttatgtatc atacacatac gatttaggtg	960
acactataga ataacatcca ctttgccctt ctctccacag gtgtccactc ccaggtccaa	1020
ctgcacctcg gttctatcga aaacgcgtcc accgtcgacc cgggcggccg cttcccttta	1080
gtgaggggta atgcttcgag cagacatgat aagatacatt gatgagtttg gacaaaccac	1140
aactagaatg cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt	1200
tgtaaccatt ataagctgca ataaacaagt taacaacaac aattgattc attttatggt	1260
tcaggttcag ggggatgtgt gggaggtttt ttaaagcaag taaaacctct acaaatgtgg	1320
taaaatccga taaggatcga tccgggctgg cgtaatagcg aagaggcccg caccgatcgc	1380
ccttcccaac agttgcgcag cctgaatggc gaatggacgc gccctgtagc ggcgcattaa	1440
gcgcggcggg tgtgggtggt acgcgcagcg tgaccgctac acttgccagc gccctagcgc	1500
ccgctccttt cgctttcttc ccttcctttc tcgccacggt cgccggcttt ccccgtaag	1560
ctctaaatcg ggggctccct ttagggttcc gatttagagc tttacggcac ctgcaccgca	1620

ES 2 784 926 T3

aaaaacttga tttgggtgat ggttcacgat cttcaatatt ggccattagc catattattc 1680  
 attggttata tagcataaat caatattggc tattggccat tgcatacggt gtatctatat 1740  
 cataatatgt acatttatat tggctcatgt ccaatatgac cgccatgttg gcattgatta 1800  
 ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc atatatggag 1860  
 ttccgcgta cataacttac ggtaaatggc ccgcctggct gaccgcccaa cgacccccgc 1920  
 ccattgacgt caataatgac gtatgttccc atagtaacgc caataggac tttccattga 1980  
 cgtcaatggg tggagtattt acggtaaact gccacttgg cagtacatca agtgtatcat 2040  
 atgccaahta cgccccctat tgacgtcaat gacggtaaat ggcccgcctg gcattatgcc 2100  
 cagtacatga ccttatggga ctttctact tggcagtaca tctacgtatt agtcatcgct 2160  
 attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg gtttgactca 2220  
 cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg gcacaaaat 2280  
 caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat gggcggtagg 2340  
 cgtgtacggt gggaggctta tataagcaga gctcgtttag tgaaccgtca gatcgctgg 2400  
 agacgccatc cacgctgttt tgacctccat agaagacacc gggaccgatc cagcctccgc 2460  
 ggccgggaac ggtgcattgg aacgcggatt cccctgcca agagtgacgt aagtaccgcc 2520  
 tatagagtct ataggccac ccccttggct tcgtagaac gcggctaaa ttaatacata 2580  
 acctatgta tcatacat acgatttagg tgacactata gaataacatc cactttgcct 2640  
 ttctctccac aggtgtccac tcccaggtcc aactgcacct cggttctatc gcgattgaat 2700  
 taattccccg gggatcctct agggtgaccg tttaaaacac cggtgccacc ggcgcgtggt 2760  
 acctctagag tcgaccggg cgccgccttc cctttagtga gggttaatgc ttcgagcaga 2820  
 catgataaga tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaatg 2880  
 ctttatttgt gaaatttgg atgctattgc tttatttga accattataa gctgcaataa 2940  
 acaagttaac aacaacaatt gcattcattt tatgtttcag gttcaggggg agatgtggga 3000  
 ggttttttaa agcaagtaaa acctctaaa atgtggtaaa atccgataag gatcgatccg 3060  
 ggctggcgta atagcgaaga ggccgcacc gatcgccctt cccaacagtt gcgcagcctg 3120  
 aatggcgaat ggacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc 3180  
 gcagcgtgac cgctacactt gccagcggc tagcggcgc tcctttcgct ttctccctt 3240  
 ctttctcgc cacgttcgccc ggctttccc gtcaagctct aaatcggggg ctcccttag 3300  
 ggttccgatt tagagcttta cggcacctcg accgcaaaa acttgatttg ggtgatggtt 3360  
 cacgtagtgg gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt 3420  
 tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt 3480

ES 2 784 926 T3

cttttgattt ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt 3540  
aacaatatatt taacgcgaat ttttaacaaaa tattaacggtt tacaatttcg cctgatgcgg 3600  
tattttctcc ttacgcatct gtgcggtatt tcacaccgca tacgcggatc tgcgcagcac 3660  
catggcctga aataacctct gaaagaggaa cttggttagg taccttctga ggcggaaaga 3720  
accagctgtg gaatgtgtgt cagttaggggt gtggaagtc cccaggtcc ccagcaggca 3780  
gaagtatgca aagcatgcat ctcaattagt cagcaaccag gtgtgaaag tccccaggct 3840  
ccccagcagc cagaagtatg caaagcatgc atctcaatta gtcagcaacc atagtcccgc 3900  
ccctaactcc gcccatcccg ccctaactc cgcccagttc cgcccattct ccgcccctg 3960  
gctgactaat ttttttatt tatgcagagg ccgaggccgc ctcggcctct gagctattcc 4020  
agaagtagtg aggaggcttt tttggaggcc taggcctttg caaaaagctt gattcttctg 4080  
acacaacagt ctggaactta aggctagagc caccatgatt gaacaagatg gattgcacgc 4140  
aggttctccg gccgcttggg tggagaggct attcggctat gactgggcac aacagacaat 4200  
cgctgctct gatgccccg tgttccggct gtcagcgag gggcgcccg ttctttttgt 4260  
caagaccgac ctgtccggtg ccctgaatga actgcaggac gaggcagcgc ggctatcgtg 4320  
gctggccacg acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag 4380  
ggactggctg ctattgggcg aagtgccggg gcaggatctc ctgtcatctc accttgetcc 4440  
tgccgagaaa gtatccatca tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc 4500  
tacctgccca ttcgaccacc aagcgaaca tcgcatcgag cgagcacgta ctcggatgga 4560  
agccggtctt gtcgatcagc atgatctgga cgaagagcat caggggctcg cgccagccga 4620  
actgttcgcc aggetcaagg cgcgcacgcc cgacggcgag gatctcgtcg tgaccatgg 4680  
cgatgcctgc ttgcccgaata tcatggtgga aaatggccgc ttttctggat tcatcgactg 4740  
tggccggctg ggtgtggcgg accgctatca ggacatagcg ttggctaccg gtgatattgc 4800  
tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg ctttacggta tcgccctcc 4860  
cgattcgag cgcatcgcct tctatcgcct tcttgaagag ttcttctgag cgggactctg 4920  
gggttcgaaa tgaccgacca agcgcagccc aacctgccat cacgatggcc gcaataaaat 4980  
atctttattt tcattacatc tgtgtgttgg ttttttgtgt gaatcgatag cgataaggat 5040  
cccggtatgg tgcaactctca gtacaactctg ctctgatgcc gcatagttaa gccagccccg 5100  
acaccgccca acaccgctg acgcgccctg acgggcttgt ctgctcccgg catccgetta 5160  
cagacaagct gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc 5220  
gaaacgcgcg agacgaaag gcctcgtgat acgcctattt ttataggta atgtcatgat 5280  
aataatggtt tcttagacgt caggtggcac ttttcgggga aatgtgcgcg gaaccctat 5340  
ttgtttattt ttctaaatac attcaaatat gtatccgctc atgagacaat aaccctgata 5400

ES 2 784 926 T3

aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgcct 5460  
tattcccttt tttgcggcat tttgccttcc tgtttttgct caccagaaa cgctggtgaa 5520  
agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa 5580  
cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt 5640  
taaagttctg ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg 5700  
tcgccgcata cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca 5760  
tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca tgagtataa 5820  
cactgcccgc aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt 5880  
gcacaacatg ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc 5940  
cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa 6000  
actattaact ggcgaaactac ttactctagc ttcccggcaa caattaatag actggatgga 6060  
ggcggataaa gttgcaggac cacttctgcg ctcggccctt ccggctggct ggtttattgc 6120  
tgataaatct ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga 6180  
tggttaagccc tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga 6240  
acgaaataga cagatcgtg agataggtgc ctactgatt aagcattggt aactgtcaga 6300  
ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat ttaaaggat 6360  
ctaggtgaag atcctttttg ataatctcat gaccaaaatc ccttaacgtg agttttcgtt 6420  
ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct 6480  
gcgcgtaate tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg tttgtttgcc 6540  
ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc 6600  
aaatactgtc cttctagtgt agccgtagt aggccaccac ttcaagaact ctgtagcacc 6660  
gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc 6720  
gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg 6780  
aacggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata 6840  
cctacagcgt gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta 6900  
tccggtgaag ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc 6960  
ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg 7020  
atgctcgtca gggggcgga gcctatgaa aaacgccagc aacgcggcct ttttacggtt 7080  
cctggccttt tgctggcctt ttgctcacat ggctcgacag atccatttaa attttaccg 7140  
tcatcaccga aacgcgcgag gcagctgtgg aatgtgtgtc agttagggtg tggaaagtcc 7200  
ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc agcaaccagg 7260

ES 2 784 926 T3

tgtggaaagt cccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag 7320  
 tcagcaacca tagtcccgcc cctaactccg cccatcccgc ccctaactcc gccagttcc 7380  
 gccattctc cgccccatgg ctgactaatt ttttttattt atgcagaggc cgaggccgcc 7440  
 tcggcctctg agctattcca gaagtagtga ggaggctttt ttggaggcct aggcttttgc 7500  
 aaaaagctaa ttcgagctcg gtacccccaa acttgacggc aatcctagcg tgaaggctgg 7560  
 taggatttta tccccgctgc catcatgggt cgaccattga actgcatcgt cgccgtgtcc 7620  
 caaaatagtg ggattggcaa gaacggagac cgaccctggc ctccgctcag gaacgagttc 7680  
 aagtacttcc aaagaatgac cacaaacctc tcagtgggaag gtaaacagaa tctgggtgatt 7740  
 atgggtagga aaacctgggt ctccattcct gagaagaatc gacctttaa ggacagaatt 7800  
 aatatagttc tcagtagaga actcaaagaa ccaccacgag gagctcattt tcttgccaaa 7860  
 agtttgatg atgccttaag acttattgaa caaccggaat tggcaagtaa agtagacatg 7920  
 gtttgatag tcggaggcag ttctgtttac caggaagcca tgaatcaacc aggccacctc 7980  
 agactccttg tgacaaggat catgcaggaa tttgaaagt acacgttttt cccagaaatt 8040  
 gatttgggga aatataaact tctcccagaa taccaggcgc tcctctctga ggtccaggag 8100  
 gaaaaaggca tcaagtataa gtttgaagtc tacgagaaga aagactaaca ggaagatgct 8160  
 ttcaagttct ctgctcccct cctaaagcta tgcattttta taagaccatg ggggatgctc 8220  
 gatcccctcg cgagttgggt cagctgctgc ctgaggctgg acgacctcgc ggagttctac 8280  
 cggcagtgca aatccgtcgg catccaggaa accagcagcg gctatccgcg catccatgcc 8340  
 cccgaactgc aggagtgggg aggcacgatg gccgctttgg tccggatcct tgtgaaggaa 8400  
 ccttacttct gtggtgtgac ataattggac aaactaccta cagagattta aagctctaag 8460  
 gtaaatataa aatttttaag tgtataatgt gttaaactac tgattctaata tgtttgtgta 8520  
 ttttagattc caacctatgg aactgatgaa tgggagcagt ggtggaatgc ctttaatgag 8580  
 gaaaacctgt tttgctcaga agaaatgcc a tctagtgatg atgaggctac tgctgactct 8640  
 caacattcta ctctccaaa aaagaagaga aaggtagaag accccaagga ctttccttca 8700  
 gaattgctaa gttttttgag tcatgctgtg tttagtaata gaactcttgc ttgctttgct 8760  
 atttacacca caaaggaaaa agctgcactg ctatacaaga aaattatgga aaaatattct 8820  
 gtaaccttta taagtaggca taacagttat aatcataaca tactgttttt tcttactcca 8880  
 cacaggcata gagtgtctgc tattaataac tatgctcaaa aattgtgtac ctttagcttt 8940  
 ttaatttgta aagggttaa taaggaatat ttgatgtata gtgccttgac tagagatcat 9000  
 aatcagccat accacatttg tagaggtttt acttgcttta aaaaacctcc cacacctccc 9060  
 cctgaacctg aaacataaaa tgaatgcaat tgttgtgtt aacttgttta ttgcagetta 9120  
 taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat ttttttcaact 9180  
 gcattctagt tgtggtttga attcggatct 9210

**REIVINDICACIONES**

1. Un ácido nucleico de vector linealizado adecuado para expresar al menos un polipéptido de interés en una célula de mamífero, que comprende
- 5 (a) al menos un casete de expresión (POI) adecuado para expresar un polipéptido de interés, donde dicho casete de expresión comprende el polinucleótido que codifica el polipéptido de interés y comprende al menos un promotor o elemento promotor/potenciador;
- 10 (b) un casete de expresión (MSM) que comprende un gen marcador seleccionable de mamífero, donde dicho gen marcador seleccionable de mamífero es un gen de resistencia a antibiótico y donde dicho casete de expresión comprende al menos un promotor o elemento promotor/potenciador;
- 15 (c) un casete de expresión (MASM) que comprende un gen marcador seleccionable, amplificable de mamífero, donde dicho casete de expresión (MASM) comprende al menos un promotor o elemento promotor/potenciador; y
- (d) un casete de expresión (PSM) que comprende un gen marcador seleccionable procariótico;
- 20 donde el casete de expresión (POI) está flanqueado en 5' por el casete de expresión (MASM), el casete de expresión (MSM) está situado en 3' respecto al casete de expresión (POI) y el casete de expresión (PSM) está situado en 3' del casete de expresión (MSM), y donde los casetes de expresión (MASM), (POI) y (MSM) están dispuestos en la misma orientación 5' a 3'.
2. El ácido nucleico de vector linealizado de acuerdo con la reivindicación 1, donde el casete de expresión (POI) comprende el polinucleótido que codifica el polipéptido de interés.
- 25 3. El ácido nucleico de vector linealizado de acuerdo con la reivindicación 1 o 2, que tiene una o más de las siguientes características:
- 30 i) dicho casete de expresión (MASM) comprende un gen que codifica una dihidrofolato reductasa (DHFR) enzimáticamente funcional;
- ii) dicho casete de expresión (MSM) comprende un gen que codifica una neomicina fosfotransferasa enzimáticamente funcional y donde dicho casete de expresión (MASM) comprende un gen que codifica una dihidrofolato reductasa (DHFR) enzimáticamente funcional;
- 35 iii) dichos casetes de expresión (POI), (MSM) y (MASM) comprenden un promotor y/o sitio de terminación de la transcripción;
- 40 iv) el casete de expresión para expresar el polipéptido de interés (POI) comprende un promotor y/o potenciador más potente que los casetes de expresión para expresar los marcadores seleccionables de mamífero;
- v) dichos casetes de expresión (POI), (MSM) y (MASM) comprenden un potenciador y/o un intrón;
- 45 vi) el casete de expresión (MASM) comprende un intrón que está situado en 3' del gen marcador seleccionable amplificable de mamífero.
4. El ácido nucleico de vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 3, donde el vector comprende al menos un casete de expresión adicional (POI') para expresar un polipéptido de interés adicional y donde el casete de expresión adicional (POI') está situado entre el casete de expresión (POI) y el casete de expresión (MSM) y donde el casete de expresión (POI') está dispuesto la misma orientación 5' a 3' que los casetes de expresión (POI) y (MSM).
- 50 5. El ácido nucleico de vector linealizado de acuerdo con la reivindicación 4 para expresar una molécula de inmunoglobulina que comprende en cada casete de expresión (POI) y (POI') un polinucleótido que codifica ya sea una cadena ligera o una pesada de una molécula de inmunoglobulina o fragmentos funcionales de esta, donde cada casete de expresión (POI) y (POI') comprende uno de dichos polinucleótidos.
- 55 6. El ácido nucleico de vector linealizado de acuerdo con la reivindicación 4 o 5, donde el casete de expresión (POI) y/o el casete de expresión (POI') si está presente comprende un promotor/potenciador de CMV y/o donde los casetes de expresión (MSM) y (MASM) comprenden un promotor/potenciador de SV40 o un promotor de SV40.
- 60

7. El ácido nucleico de vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 6, donde el ácido nucleico de vector tiene una o más de las siguientes características
- 5 i) el gen marcador seleccionable procariótico proporciona una resistencia a antibiótico, donde dicho antibiótico se selecciona del grupo que consiste en ampicilina, kanamicina, tetraciclina y cloranfenicol;
- ii) el casete de expresión (PSM) está dispuesto en la misma orientación 5' a 3' que los casetes de expresión (POI), (MSM) y (MASM).
- 10 8. El ácido nucleico de vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 7, donde el vector de expresión se selecciona del grupo que consiste en
- (a) un ácido nucleico de vector lineal que comprende los siguientes elementos genéticos en la disposición indicada, donde la dirección 5' a 3' se indica mediante la →:
- 15 I. Promotor del casete de expresión (MASM) (→)
- II. Gen que codifica el marcador seleccionable amplificable de mamífero del casete de expresión (MASM) (→)
- 20 III. Intrón del casete de expresión (MASM) (→)
- IV. Sitio de poliA del casete de expresión (MASM) (→)
- V. Promotor del casete de expresión (POI) (→)
- 25 VI. Intrón del casete de expresión (POI) (→)
- VII. Polinucleótido que codifica un polipéptido de interés, que se inserta en el casete de expresión (POI) (→)
- 30 VIII. Sitio de poliA del casete de expresión (POI) (→)
- IX. Promotor del casete de expresión (POI') (→)
- X. Intrón del casete de expresión (POI') (→)
- 35 XI. Polinucleótido que codifica un polipéptido de interés adicional, que se inserta en el casete de expresión (POI') (→)
- XII. Sitio de poliA del casete de expresión (POI') (→)
- 40 XIII. Promotor del casete de expresión (MSM) (→)
- XIV. Gen que codifica el marcador seleccionable de mamífero del casete de expresión (MSM) (→)
- 45 XV. Sitio de poliA del casete de expresión (MSM) (→)
- XVI. Casete de expresión PSM (→) o (←);
- (b) un ácido nucleico de vector como el que se muestra como la Seq. ID No. 1 o Seq. ID No. 16 o un derivado de estas, que comprende la misma disposición de elementos genéticos.
- 50 9. Una célula hospedadora de mamífero que comprende un ácido nucleico de vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 8.
- 55 10. La célula hospedadora de mamífero de acuerdo con la reivindicación 9, donde el ácido nucleico de vector linealizado está integrado de forma estable en el genoma de la célula hospedadora de mamífero.
11. Un método para producir una célula hospedadora de mamífero que expresa un polipéptido de interés, que comprende transfectar de forma estable una célula hospedadora de mamífero con el ácido nucleico de vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 8.
- 60

- 5 12. El método de la reivindicación 11, que comprende linealizar un vector circular antes de la transfección para proporcionar el ácido nucleico de vector linealizado de acuerdo con al menos una de las reivindicaciones 1 a 8, donde el vector circular comprende un sitio de restricción para la linealización predefinido, preferentemente único, que está situado en la forma circular del vector entre los casetes de expresión (PSM) y (MASM).
- 10 13. Un método para producir un polipéptido de interés, comprendiendo dicho método cultivar al menos una célula hospedadora de mamífero de acuerdo con la reivindicación 9 o 10 o una célula hospedadora de mamífero obtenida de acuerdo con la reivindicación 11 o 12 en un medio de cultivo celular en condiciones que permitan la expresión de dicho polipéptido de interés.
- 15 14. El método de acuerdo con la reivindicación 13, que tiene una o más de las siguientes características:
- i) el método comprende
- 15 - aislar el polipéptido de interés a partir de dicho medio de cultivo celular y/o a partir de dicha célula hospedadora y/o
- procesar el polipéptido de interés aislado;
- 20 ii) dicho polipéptido de interés se secreta en el medio de cultivo celular y el método comprende aislar el polipéptido de interés a partir del medio de cultivo celular.



Fig. 2

